

Db 203 GPDDRLYTFCCKLPTKFSMKHLDLNCST-SSVPVVKVMVVEECNCE 248

RESULT 14

O35793 PRELIMINARY; PRT; 184 AA.

AC O35793; DT 01-JAN-1998 (TRENBLrel. 05, Created)

DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)

DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)

DE DRM protein.

GN DRM.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=97378055; PubMed=9234736;

RA Topol L.Z., Marx M., Laugier D., Bogdanova N.N., Boubnov N.V., Clausen P.A., Calothy G., Blair D.G.;

RT "Identification of drm, a novel gene whose expression is suppressed in transformed cells and which can inhibit growth of normal but not transformed cells in culture.";

RL Mol. Cell. Biol. 17:4801-4810(1997).

DR EMBL; Y10019; CAA71126.1; -.

DR InterPro; IPR000359; Cys knot.

DR Pfam; PF03045; DAN_dom.

DR SMART; SM00041; DAN; 1.

SQ SEQUENCE 184 AA; 20680 MW; 90AC99749B1ADB8F6 CRC64;

Query Match 8.3%; Score 86; DB 11; Length 184;

Best Local Similarity 23.3%; Pred. No. 0.46;

Matches 37; Conservative 24; Mismatches 60; Indels 38; Gaps 9;

QY 4 QAFRNDATVPIGLGEYPEPPENNTQWRAENGRRPHHPYDAKDVSYSCRELHYT-- 61

Db 38 KAQNDSEQT-----QSPFGSRTGRGQ--GRGTAMP--GEEVLESSQEALHVT 86

QY 62 RFLTDGPCR-----SAKPVTELVCSCGCGPARLLPNAIGRVKWRPNPDPFR 108

Db 87 KYLRDWCKTQPLKQTIHEEGCNSRTIINRFYGCQN-SFYIPRH-----RKEGSFQ 139

QY 109 ----CIPDRYRAQRVQLCPGGAAP-RSRKRVLVASCKC 142

Db 140 SCSFCKPKKFTTMMVTLNCPPELPPTTKKRVTRVKQCRC 178

Search completed: March 23, 2003, 14:18:47

Job time : 54.28 secs

RT "Cloning of the murine DRM gene and characterization of its oncogene suppressible promoter.";

RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RC TISSUE=COLON;

RA Strausberg R.;

RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF045801; AAC40111.1; -.

DR EMBL; AF108189; AAD54056.1; -.

DR EMBL; BC015293; AAH15293.1; -.

DR MGD; MGI:1344337; Cktsf1b1.

DR InterPro; IPR000359; Cys knot.

DR InterPro; IPR004133; DAN_dom.

DR Pfam; PF03045; DAN; 1.

DR SMART; SM00041; CT; 1.

SQ SEQUENCE 184 AA; 20710 MW; D1BD99783CDC0F2D CRC64;

Query Match 8.3%; Score 86; DB 11; Length 184;

Best Local Similarity 23.3%; Pred. No. 0.46;

Matches 37; Conservative 24; Mismatches 60; Indels 38; Gaps 9;

QY 4 QAFRNDATVPIGLGEYPEPPENNTQWRAENGRRPHHPYDAKDVSYSCRELHYT-- 61

Db 38 KAQNDSEQT-----QSPFGSRTGRGQ--GRGTAMP--GEEVLESSQEALHVT 86

QY 62 RFLTDGPCR-----SAKPVTELVCSCGCGPARLLPNAIGRVKWRPNPDPFR 108

Db 87 KYLRDWCKTQPLKQTIHEEGCNSRTIINRFYGCQN-SFYIPRH-----RKEGSFQ 139

QY 109 ----CIPDRYRAQRVQLCPGGAAP-RSRKRVLVASCKC 142

Db 140 SCSFCKPKKFTTMMVTLNCPPELPPTTKKRVTRVKQCRC 178

Search completed: March 23, 2003, 14:18:47

Job time : 54.28 secs

Cytogenet. Cell Genet. 89:79-84 (2000).

RL EMBL; AF045800; AAC39725.1; -

DR EMBL; AB032372; BAA84462.1; -

DR EMBL; AF110137; AAF06677.1; -

DR EMBL; AF154054; AAG23891.1; -

DR InterPro; IPR000359; Cys knot.

DR InterPro; IPR004133; DAN_dom.

DR Pfam; PF03045; DAN; 1.

DR SMART; SM00041; CT; 1.

SQ SEQUENCE 184 AA; 20697 MW; 4B588598DEL2C47E CRC64;

Query Match 8.6%; Score 88.5; DB 4; Length 184;

Best Local Similarity 23.4%; Pred. No. 0.25;

Matches 36; Conservative 20; Mismatches 55; Indels 43; Gaps 8;

Qy 23 PPNENQ-----TNRAENGGRPPHPHYDAKDVSEYSCRELHYT--RFLTD 66

Db 34 PPPDKACHNDSEQTQSPQPGSRNRGRGQGTAMP--GEVLESSQEAHVTERKYLK 91

Qy 67 GPCR-----SAKPVTELVCSGCGCPARLLPNAIGRVKWRPNPDPFR---C 109

Db 92 DWCKTQPLKQTHIEBGCNSRTIINRFYCGQN-SFYIPRHI-----RKEEGSFQSCSFC 144

Qy 110 IPDRYRAORVQLCPGGAAP-RSRKVRIVASCKC 142

Db 145 KPKKFTTMMVTILNCPQLPPTKKKRVTRVKQCR 178

RESULT 11

Q8WNY1 PRELIMINARY; PRT; 184 AA.

AC Q8WNY1;

DT 01-MAR-2002 (TrEMBLrel. 20, Created)

DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)

DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)

DE Gremlin.

OS Macaca mulatta (Rhesus macaque).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;

OC Cercopitheciinae; Macaca.

OX NCBI_TaxID=9544;

RN [1]

RP SEQUENCE FROM N.A.

RA Christenson L.K., Duffy D.M.;

RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF439783; AAL32022.1; -

DR InterPro; IPR000359; Cys knot.

DR InterPro; IPR004133; DAN_dom.

DR Pfam; PF03045; DAN; 1.

DR SMART; SM00041; CT; 1.

SQ SEQUENCE 184 AA; 20697 MW; 4B588598DEL2C47E CRC64;

Query Match 8.6%; Score 88.5; DB 6; Length 184;

Best Local Similarity 23.4%; Pred. No. 0.25;

Matches 36; Conservative 20; Mismatches 55; Indels 43; Gaps 8;

Qy 23 PPNENQ-----TNRAENGGRPPHPHYDAKDVSEYSCRELHYT--RFLTD 66

Db 34 PPPDKACHNDSEQTQSPQPGSRNRGRGQGTAMP--GEVLESSQEAHVTERKYLK 91

Qy 67 GPCR-----SAKPVTELVCSGCGCPARLLPNAIGRVKWRPNPDPFR---C 109

Db 92 DWCKTQPLKQTHIEBGCNSRTIINRFYCGQN-SFYIPRHI-----RKEEGSFQSCSFC 144

Qy 110 IPDRYRAORVQLCPGGAAP-RSRKVRIVASCKC 142

Db 145 KPKKFTTMMVTILNCPQLPPTKKKRVTRVKQCR 178

RESULT 12

Q40600 PRELIMINARY; PRT; 332 AA.

ID Q40600

AC Q40600;

DT 01-NOV-1996 (TrEMBLrel. 01, Created)

DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)

DE 70S mitochondrial ribosomal protein L2.

GN RP12.

OS Oenothera lutea (Bertero's evening primrose).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurosids II; Myrtales; Onagraceae; Oenothera.

OX NCBI_TaxID=3950;

RN [1]

RP SEQUENCE FROM N.A.

RA Fritz I., Schuster W.;

RL Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.

DR EMBL; X80170; CAA56451.1; -

DR InterPro; IPR002171; Ribosomal L2.

DR Pfam; PF00181; Ribosomal_L2; 1.

DR Ribosomal protein.

KW Ribosomal protein.

SQ SEQUENCE 332 AA; 36678 MW; B8E123914F01791B CRC64;

Query Match 8.5%; Score 88; DB 10; Length 332;

Best Local Similarity 25.5%; Pred. No. 0.54;

Matches 27; Conservative 16; Mismatches 35; Indels 28; Gaps 5;

Qy 3 WOAFNRDATEVIPCGLGPEPPPPNNQTNRAENGGRPPHPHYDAKDVSEYSCRELHYT 62

Db 244 WLSFRQETDGLVGAAEHNSKPKTDQ-----GSLP-----AKPIGEGT----- 282

Qy 63 FLTDPGCR-SAKPVTELVCSGCGCPARLLPNAIGRVKWRPNPDPF 107

Db 283 --KDGACKVDRAPTYIIASHQLEAGKVMVNC---DWSKPSSTDF 322

RESULT 13

Q9PWB0 PRELIMINARY; PRT; 272 AA.

ID Q9PWB0;

AC Q9PWB0;

DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)

DE Cerberus homolog.

GN CER.

OS Gallus gallus (Chicken).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;

OC Gallus.

OX NCBI_TaxID=9031;

RN [1]

RP SEQUENCE FROM N.A.

RP MEDLINE=99439862; PubMed=10508582;

RA Zhu L., Marvin M.J., Gardiner A., Lassar A.B., Mercola M., Stern C.D.,

RA Levin M.;

RT "Cerberus regulates left-right asymmetry of the embryonic head and

RT heart.";

RL Curr. Biol. 9:931-938 (1999).

DR EMBL; AF139721; AAD51610.1; -

DR InterPro; IPR000359; Cys knot.

DR InterPro; IPR004133; DAN_dom.

DR Pfam; PF03045; DAN; 1.

DR SMART; SM00041; CT; 1.

DR PROSITE; PS01225; CTCK 2; 1.

SQ SEQUENCE 272 AA; 31224 MW; 36E4C9F719711BCA CRC64;

Query Match 8.4%; Score 87; DB 13; Length 272;

Best Local Similarity 21.5%; Pred. No. 0.55;

Matches 23; Conservative 20; Mismatches 42; Indels 22; Gaps 4;

Qy 44 PYDAKDVSEYSCRELHYTFLTPGCRSAKPVTELVCSGCGCPARLLPNAIGRVKWRPN 103

Db 157 PIKTNMHQETCTRLPFSQVAHESCEKV-IVQNNLCFGKCS-----FHPV 202

Qy 104 GPDER-----CIPDRYRAORVQLCPGGAAPRSKRVRIVASCKCK 143

GN MUC5AC.
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9825;
 RN [1]
 RP SEQUENCE OF 35-109 FROM N.A.
 RC TISSUE=GASTRIC EPITHELIUM;
 RA MEDLINE=95275264; PubMed=7755593;
 RX Turner B.S., Bhaskar K.R., Hadzopoulou-Cladaras M., Specian R.D.,
 RA LaMont J.T.;
 RT "Isolation and characterization of cDNA clones encoding pig gastric
 RT mucin.";
 RL Biochem. J. 308:89-96 (1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=GASTRIC EPITHELIUM;
 RX MEDLINE=99431563;
 RA Turner B.S., Bhaskar K.R., Hadzopoulou-Cladaras M., LaMont J.T.;
 RT "Cysteine-rich regions of pig gastric mucin contain von willebrand
 RT factor and cysteine knot domains at the carboxyl terminal (1).";
 RL Biochim. Biophys. Acta 1447:77-92 (1999).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=GASTRIC EPITHELIUM;
 RA Turner B.S., Bhaskar K.R., Hadzopoulou-Cladaras M., LaMont J.T.;
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF054584; AAD19833.1; -
 DR InterPro; IPR000359; Cys_knot.
 DR InterPro; IPR002400; GF_Cysknot.
 DR InterPro; IPR001007; VWF_C.
 DR InterPro; IPR001846; VWF_D.
 DR Pfam; PF00007; Cys_knot; 1.
 DR Pfam; PF00093; vwc; 1.
 DR Pfam; PF00094; vwd; 1.
 DR PRINTS; PR00438; GFCYSKNOT.
 DR SMART; SM00041; CT; 1.
 DR SMART; SM00214; VWC; 2.
 DR SMART; SM00216; VWD; 1.
 DR PROSITE; PS01185; CTCK_1; UNKNOWN_1.
 DR PROSITE; PS01225; CTCK_2; 1.
 DR PROSITE; PS01208; VWF; UNKNOWN_2.
 FT NON TER 1
 SQ SEQUENCE 988 AA; 105264 MW; 786F949E0FCE8465 CRC64;
 Query Match 8.7%; Score 90; DB 6; Length 988;
 Best Local Similarity 26.4%; Pred. No. 1.1;
 Matches 37; Conservative 16; Mismatches 55; Indels 32; Gaps 7;
 Qy 51 SEYSCRELHYTRFLTDGPCRSAPVTELVCSGCGPARLL--PNAIGRVKWRPNGPDPR 108
 Db 861 NOSSCAVYHQHVLOQSCSAGFRLTYCGNCGDTASWYSPEA-----NAVEHR 911
 Qy 109 ---CIPDRYRAQRVQLLCPGGAAPRSRKRLVASCKC--KRLTRFHNSQLKDFGPETAR 163
 Db 912 CKCCQELQVALRVNVLHCPDGSS-RAFSYTEVEKCCGVCQRCDSHGDLSLSEEPQLSR 970
 Qy 164 -----POKGRKRP 172
 Db 971 DAGHGLWRTGAPO-----PRP 986
 RESULT 9
 O73754
 ID O73754 PRELIMINARY; PRT; 182 AA.
 AC O73754;
 DT 01-AUG-1998 (TREMBLrel. 07, Created)
 DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Gremlin.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae;
 RT expressed in adult and fetal brain.";

OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=OVARY;
 RX MEDLINE=98325381; PubMed=9660951;
 RA Hsu D.R., Economides A.N., Wang X., Eimon P.M., Harland R.M.;
 RT "The Xenopus dorsaling factor Gremlin identifies a new family of
 RT secreted proteins that antagonize BMP activities.";
 RL Mol. Cell 1:673-683 (1998).
 RL EMBL; AF045798; AAC41279.1; -
 DR InterPro; IPR000359; Cys_knot.
 DR InterPro; IPR004133; DAN_dom.
 DR Pfam; PF03045; DAN; 1.
 DR SMART; SM00041; CT; 1.
 SQ SEQUENCE 182 AA; 20374 MW; 4EF3A655FFA1B7CE CRC64;
 Query Match 8.6%; Score 88.5; DB 13; Length 182;
 Best Local Similarity 23.5%; Pred. No. 0.25;
 Matches 35; Conservative 23; Mismatches 54; Indels 37; Gaps 8;
 Qy 23 PPENNTQNMRAENG-----GRPHHPYDAKOVSEYSCRELHYT--RFLTDGPCR- 70
 Db 36 PPPDKGP-NDSEGGQAPQDGVKGGKQALAAEVLSSQBALHYTERKYLKRDWCKT 94
 Qy 71 -----SAKPVTELVCSGCGPARLLPNAIGRVKWRPNGPDPR----CIPDRY 114
 Db 95 OPLKQTIHEDGCSRNTIINRFYGCQN-SFYIPRHIRREE-----GSFQSCSFCXPKKF 147
 Qy 115 RAQRVQLLCPGGAAP-RSRKRLVASCKC 142
 Db 148 TTMVTLNCPQLQPPTKKKRITRVKQCRC 176
 RESULT 10
 O60565
 ID O60565 PRELIMINARY; PRT; 184 AA.
 AC O60565;
 DT 01-AUG-1998 (TREMBLrel. 07, Created)
 DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE GREMLIN (DRM).
 GN DRM.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98325381; PubMed=9660951;
 RA Hsu D.R., Economides A.N., Wang X., Eimon P.M., Harland R.M.;
 RT "The Xenopus dorsaling factor Gremlin identifies a novel family of
 RT secreted proteins that antagonize BMP activities.";
 RL Mol. Cell 1:673-683 (1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Tate G., Mitsuya T.;
 RT "Human Gremlin homologue.";
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA McMahon R.A., Murphy M., Clarkson M.R., Godson C., Martin F.,
 RA Brady H.R.;
 RT "IHG-2, a mesangial cell gene differentially induced in high glucose,
 RT is human gremlin.";
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC TISSUE=SMALL INTESTINES;
 RX MEDLINE=20354996; PubMed=10894942;
 RA Topol L.Z., Modi W.S., Koochekpour S., Blair D.G.;
 RT "DRM/GREMLIN (CKTSFIB1) maps to human chromosome 15 and is highly
 RT expressed in adult and fetal brain.";

```
DR InterPro: IPR000359; Cys knot.
DR InterPro: IPR004133; DAN_dom.
DR Pfam: PF03045; DAN; 1.
DR SMART; SM00041; CT; 1.
SQ SEQUENCE 184 AA; 21166 MW; 3510B44E8F6D5EA CRC64;

Query Match      8.8%; Score 90.5; DB 13; Length 184;
Best Local Similarity 22.9%; Pred. No. 0.16;
Matches 35; Conservative 23; Mismatches 60; Indels 35; Gaps 8;

QY 18 GEYEP-----PPENQTMNRAENGRRPHI-----PYDAKDVSEYSCRELHYT--RFLTDG 67
Db 33 GAIPPPKQDQNDSEQQTQSGSRHRRERKGTSPABEVLESSQEAUHTIRKILKRD 92
QY 68 PCR-----SAKPVTELVCSCGCGPARLLPNAIGRVKWRPNPGRPPFR-----CI 110
Db 93 WCKTQPLKQTHIEGSGNSRIINFCYGCQN-SFYIPRHV-----RKEGSGFQSCSFK 145
QY 111 PDYRAQRVQLLCFPGAAPRSRK-VRLVASC 142
Db 146 PKFTTMTVTLNCPQLPQPPRKKKRIIVKRCRC 178

RESULT 5
QY557 ID QY557 PRELIMINARY; PRT; 712 AA.
AC QY557;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE DJ1163J1.3 (Similar to mouse B99 protein) (Fragment).
GN DJ1163J1.3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Lloyd D.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
FT EMBL; AL031588; CAB38415.1; -.
FT NON_TER 1
SQ SEQUENCE 712 AA; 75685 MW; ECBD21333EF72F32 CRC64;

Query Match      8.7%; Score 90; DB 4; Length 712;
Best Local Similarity 22.7%; Pred. No. 0.78;
Matches 34; Conservative 20; Mismatches 36; Indels 60; Gaps 6;

QY 68 PCSAKPVTELVCSCGCGPARLLP-----NAIGRVKWRPNPGRPFRCIPDRYRAQRV----- 119
Db 333 PANSSRPLSNISKSGRMGPAMLRLPALPAGPVGASSW-----QAKRVDVSE 377
QY 120 ----QLLCPGGAAP-----RSRKVLVASC-----KCKELT 146
Db 378 LAAEQLTAPPASPTQPTPEGGGWLNSSCAWSESSQLNKTIRRRDSCLSNKTWKMP 437
QY 147 RFHQSELKDF-----GPETARPQKGRKRP 172
Db 438 TPTNQFKIPKFSIGDSPDSSTPKLSRAQRP 467

RESULT 6
QY523 ID QY523 PRELIMINARY; PRT; 720 AA.
AC QY523;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE B99.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;

[1]
RN SEQUENCE FROM N.A.
RP MEDLINE=20432105; PubMed=10974554;
RA Monte M., Collavin L., Lazarevic D., Utrera R., Dragani T.A.,
RA Schneider C.;
RT "Cloning, chromosome mapping and functional characterization of a
RT human homologue of murine Gtse-1 (B99) gene.";
RL Gene 254:229-236(2000).
DR EMBL; AF223408; AAF31459.1; -.
SQ SEQUENCE 720 AA; 76628 MW; D4E852CE4BE85BBF CRC64;

Query Match      8.7%; Score 90; DB 4; Length 720;
Best Local Similarity 22.7%; Pred. No. 0.79;
Matches 34; Conservative 20; Mismatches 36; Indels 60; Gaps 6;

QY 68 PCSAKPVTELVCSCGCGPARLLP-----NAIGRVKWRPNPGRPFRCIPDRYRAQRV----- 119
Db 341 PANSSRPLSNISKSGRMGPAMLRLPALPAGPVGASSW-----QAKRVDVSE 385
QY 120 ----QLLCPGGAAP-----RSRKVLVASC-----KCKELT 146
Db 386 LAAEQLTAPPASPTQPTPEGGGWLNSSCAWSESSQLNKTIRRRDSCLSNKTWKMP 445
QY 147 RFHQSELKDF-----GPETARPQKGRKRP 172
Db 446 TPTNQFKIPKFSIGDSPDSSTPKLSRAQRP 475

RESULT 7
QY580 ID QY580 PRELIMINARY; PRT; 720 AA.
AC QY580;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE G-2 and S-phase expressed 1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Straussberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC006325; AAH06325.1; -.
SQ SEQUENCE 720 AA; 76614 MW; ACD91CCCD008A89C CRC64;

Query Match      8.7%; Score 90; DB 4; Length 720;
Best Local Similarity 22.7%; Pred. No. 0.79;
Matches 34; Conservative 20; Mismatches 36; Indels 60; Gaps 6;

QY 68 PCSAKPVTELVCSCGCGPARLLP-----NAIGRVKWRPNPGRPFRCIPDRYRAQRV----- 119
Db 341 PANSSRPLSNISKSGRMGPAMLRLPALPAGPVGASSW-----QAKRVDVSE 385
QY 120 ----QLLCPGGAAP-----RSRKVLVASC-----KCKELT 146
Db 386 LAAEQLTAPPASPTQPTPEGGGWLNSSCAWSESSQLNKTIRRRDSCLSNKTWKMP 445
QY 147 RFHQSELKDF-----GPETARPQKGRKRP 172
Db 446 TPTNQFKIPKFSIGDSPDSSTPKLSRAQRP 475

RESULT 8
QY786 ID QY786 PRELIMINARY; PRT; 988 AA.
AC QY786;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Gastric mucin (fragment).
```


<hr/>							
RC	TISSUE=KIDNEY;						
RA	Submitted (JAN-2002) to the EMBL/GenBank/DDBJ databases.						
DR	EMBL; AK007967; BAB25378.1; -						
DR	EMBL; AK002240; BAB21957.1; -						
DR	EMBL; AK002396; BAB22068.1; -						
DR	EMBL; AK007893; BAB25333.1; -						
DR	EMBL; BC021458; AAH21458.1; -						
DR	MGI; MGI:1913292; 0610006G05Rik.						
DR	InterPro; IPR00359; Cys_knot.						
DR	PROSITE; PS01225; CTCCK_2; 1.						
SQ	SEQUENCE 206 AA; 23174 MW; 70D24819EB06CBC9 CRC64;						
<hr/>							
	Query Match	32.0%;	Score 330;	DB 11;	Length 206;		
	Best Local Similarity	39.8%;	Pred. No. 2.8e-26;				
	Matches 74;	Conservative 35;	Mismatches 65;	Indels 12;	Gaps 7;		
<hr/>							
QY	5 AFRNDATEVIPGLGEYPEPP-PENNQTMMNRAENGRRPHHP---YDAKDVSEVSCRELHY	60					
Dd	23 AFKNDATEILYSHVVKVPVAHPSPNSSTLNQAANGRR--HFSNTGLDRNSRVQVCRELRS	80					
<hr/>							
QY	61 TRFLTDCPCSAKPVTVELVCSGCCGPBARLLPNAIG---RVKWM-RPNGDPFRCIPDRYRA	116					
Dd	81 TKYSISDQCCTISPLKELVCAGECULPVLPNWIGGGYGTKYWSRSSSQEWRCVNDKTRT	140					
<hr/>							
QY	117 QRVLQLCPGAAPRKRLVASCKKCLRFHQSLKDFGFETARP-QKGKRPFGAK	175					
Dd	141 QRIQLCCQGSGT-RYYKTVTVTACKKYTROHNESHNFESVSAPKAQHHRKRASK	199					
<hr/>							
QY	176 ANQAEI 181						
Dd	::::						
Dd	200 SSKHSL 205						
<hr/>							
RESULT 2							
Q96HJ7	ID Q96HJ7 PRELIMINARY; PRT; 206 AA.						
AC	Q96HJ7;						
DT	01-DEC-2001 (TrEMBLrel. 19, Created)						
DT	01-DEC-2001 (TrEMBLrel. 19, Last sequence update)						
DE	Similar to RIKEN cDNA 0610006G05 gene (CDA019).						
OS	Homo sapiens (Human);						
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;						
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.						
NCBI_TaxID=9606;	[1]						
RN	SEQUENCE FROM N.A.						
RP	TISSUE=BONE MARROW;						
RC	Strausberg R.;						
RL	Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.						
RN	[2]						
RP	SEQUENCE FROM N.A.						
RA	Liu F., Xu X.R., Qian B.Z., Xiao H., Chen Z., Han Z.;						
RT	"A novel gene expressed in human pheochromocytoma";						
RL	Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.						
DR	EMBL; BC008484; AAH08484.1; -						
DR	EMBL; AF361494; AAU57219.1; -						
DR	InterPro; IPR00359; Cys_knot.						
DR	PROSITE; PS01225; CTCCK_2; 1.						
SQ	SEQUENCE 206 AA; 23306 MW; 9FB3CC41B4B53834 CRC64;						
<hr/>							
	Query Match	31.8%;	Score 328;	DB 4;	Length 206;		
	Best Local Similarity	39.9%;	Pred. No. 4.5e-26;				
	Matches 73;	Conservative 36;	Mismatches 62;	Indels 12;	Gaps 7;		
<hr/>							
QY	5 AFRNDATEVIPGLGEYPEPP-PENNQTMMNRAENGRRPHHPYDAKDVS---EYSCRELHY	60					
Dd	23 AFKNDATEILYSHVVKVPVAHPSPNSSTLNQAANGRR--HFSNTGLDRNTVTVQVCRELRS	80					
<hr/>							
QY	61 TRFLTDCPCSAKPVTVELVCSGCCGPBARLLPNAIG---RVKWM-RPNGDPFRCIPDRYRA	116					
Dd	81 TKYSISDQCCTISPLKELVCAGECULPVLPNWIGGGYGTKYWSRSSSQEWRCVNDKTRT	140					

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 28, 2003, 14:12:16 ; Search time 53.28 Seconds

(without alignments)
715.441 Million cell updates/sec

Title: US-09-867-274-4

Perfect score: 1032

Sequence: 1 QGWAQFRDATEVIFGLGEY.....KGRKPRPGAKANQALENAY 185

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 21:*

1: sp_archaea:*

2: sp_bacteria:*

3: sp_fungi:*

4: sp_human:*

5: sp_invertebrate:*

6: sp_mammal:*

7: sp_mhc:*

8: sp_organelle:*

9: sp_phage:*

10: sp_plant:*

11: sp_rodent:*

12: sp_virus:*

13: sp_vertebrate:*

14: sp_unclassified:*

15: sp_rvirus:*

16: sp_bacteriap:*

17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	330	32.0	206	11 Q9CQ4	Q9CQ4 mus musculus
2	328	31.8	206	4 Q96HJ7	Q96HJ7 homo sapien
3	249	24.1	134	4 Q9X3U3	Q9X3U3 homo sapien
4	90.5	8.8	184	13 Q73755	Q73755 gallus gall
5	90	8.7	712	4 Q9Y557	Q9Y557 homo sapien
6	90	8.7	720	4 Q9NVZ3	Q9NVZ3 homo sapien
7	90	8.7	720	4 Q9BRE0	Q9BRE0 homo sapien
8	90	8.7	988	6 Q97867	Q97867 sus scrofa
9	88.5	8.6	182	13 Q73754	Q73754 xenopus lae
10	88.5	8.6	184	4 Q60565	Q60565 homo sapien
11	88.5	8.6	184	6 Q8WNY1	Q8WNY1 macaca mula
12	88	8.5	332	10 Q40600	Q40600 ctenodonta b
13	87	8.4	272	13 Q9PWB0	Q9PWB0 gallus gall
14	86	8.3	184	11 Q35793	Q35793 rattus norv
15	86	8.3	184	11 Q70326	Q70326 mus musculus
16	86	8.3	1081	4 Q76065	Q76065 homo sapien

17	85.5	8.3	215	8 Q9MFB6	Q9MFB6 beta vulgar
18	85	8.2	272	13 Q9PUK2	Q9PUK2 gallus gall
19	85	8.2	1035	5 Q9NEG1	Q9NEG1 drosophila
20	85	8.2	1114	11 Q9JKW7	Q9JKW7 mus musculus
21	84.5	8.2	1223	11 Q9QW33	Q9QW33 rattus sp.
22	83.5	8.1	279	10 Q8VZ42	Q8VZ42 arabidopsis
23	83	8.0	644	12 Q65732	Q65732 bluetongue
24	83	8.0	644	12 Q65751	Q65751 bluetongue
25	83	8.0	667	10 Q8S1G5	Q8S1G5 oryza sativ
26	82.5	8.0	141	6 Q95J85	Q95J85 monodelphis
27	82.5	8.0	168	11 Q88273	Q88273 mus musculus
28	82.5	8.0	334	10 Q49558	Q49558 arabidopsis
29	82	7.9	1637	6 Q9XSV8	Q9XSV8 bos taurus
30	82	7.9	5146	6 Q8SPM4	Q8SPM4 bos taurus
31	81.5	7.9	262	13 Q13091	Q13091 pleurodeles
32	81.5	7.9	392	10 Q9AX94	Q9AX94 oryza sativ
33	80.5	7.8	157	6 Q8WNC7	Q8WNC7 pongo pygma
34	80.5	7.8	157	6 Q8WNC5	Q8WNC5 pongo pygma
35	80.5	7.8	168	4 Q9H772	Q9H772 pongo pygma
36	80.5	7.8	270	13 Q70041	Q70041 homo sapien
37	80.5	7.8	386	10 Q8S5P2	Q8S5P2 oryza sativ
38	80.5	7.8	1704	5 Q94446	Q94446 chironomus
39	80.5	7.8	384	11 Q9CUP6	Q9CUP6 mus musculus
40	80	7.8	766	13 Q98SW2	Q98SW2 oncorhynchus
41	80	7.8	1042	4 Q13792	Q13792 homo sapien
42	80	7.7	157	6 Q8WNC6	Q8WNC6 pongo pygma
43	79.5	7.7	157	6 Q8WN18	Q8WN18 alluropoda
44	79	7.7	141	6 Q8WN18	Q8WN18 alluropoda
45	79	7.7	2026	4 Q00468	Q00468 homo sapien

ALIGNMENTS

RESULT 1

ID	Q9CQ4	PRELIMINARY;	PRT;	206 AA.
AC	Q9CQ4;			
DT	01-JUN-2001 (TREMREL. 17, Created)			
DT	01-JUN-2001 (TREMREL. 17, Last sequence update)			
DT	01-MAR-2002 (TREMREL. 20, Last annotation update)			
DE	0610006G05RIK protein (RIKEN CDNA 0610006G05 gene).			
GN	0610006G05RIK.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=CS7BL/6J; TISSUE=PANCREAS, AND KIDNEY;			
RX	MEDLINE=21085660; PubMed=11217851;			
RA	Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,			
RA	Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,			
RA	Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,			
RA	Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,			
RA	Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,			
RA	Fleischmann W., Gaasterland T., Nikaido I., Pesole G., Quackenbush J.,			
RA	Kuehl P., Lewis S., Mateu Y., Suzuki R., Tomita M., Wagner L., Washio T.,			
RA	Schriml L.M., Staubli P., Furuno M., Aono H., Baldarelli R., Barsh G.,			
RA	Sakai K., Okido T., Furuno M., Carninci P., de Bonaldo M.F.,			
RA	Blake J., Boffelli D., Bojunga N., Carninci P., Fujita M., Gariboldi M.,			
RA	Brownstein M.J., Bult C., Fletcher C., Hume D.A., Kamiya M., Lee N.H.,			
RA	Gustincich S., Hill D., Hofmann J., Mazzarelli J., Mombaerts P.,			
RA	Lyons P., Marchionni L., Mashima J., Rodriguez I., Sakamoto N.,			
RA	Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,			
RA	Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,			
RA	Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,			
RA	Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohseki S.,			
RA	Hayashizaki Y.;			
RT	"Functional annotation of a full-length mouse cDNA collection.";			
RL	Nature 409:685-690(2001).			
RN	[2]			
RP	SEQUENCE FROM N.A.			

Search completed: March 28, 2003, 14:16:52
Job time : 16.3067 secs

Mon Mar 31 09:44:37 2003

```

FT DISULFID 114 121 BY SIMILARITY.
FT CARBOHYD 34 34 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 141 AA; 15060 MW; C6CF98036B3C4EE0 CRC64;

Query Match 8.4%; Score 86.5; DB 1; Length 141;
Best Local Similarity 28.8%; Pred. No. 0.36; Indels 17; Gaps 5;
Matches 34; Conservative 9; Mismatches 59;

QY 55 CRELHYTRFLTDGCRSAKPVTELVCSCGQC-GPARLLPNAIGRVKWRPNPGRDPFCIPDR 113
DB 30 CRPTNATLAESDACPVCTVFTTICAGYCPSMVRVLPAL-----PPGQLVCTYRE 82

QY 114 YRAQRVOLL-CPGAAP-RSRKRLVASCCKELTRFHNOSELKDFGPETARPQGRKP 170
DB 83 LSFSSIRLPGCPGVDPFIFSPVALSCSCGSLSH-----SDCGPRAPRPHICTRP 134

RESULT 11
VP4_BTIV11 STANDARD; PRT; 644 AA.
AC P33428;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE VP4 core protein.
GN S4.
OS Bluetongue virus (serotype 11 / isolate USA).
OC Viruses; dsRNA viruses; Reoviridae; Orbivirus.
OX NCBI_TaxID=33716;
RN [1]
RP MEDLINE=93331736; PubMed=8393245;
RX Huang I.J., Hayama E., Jeong Y.J., Li J.K.-K.;
RT "Conservation of the segment 4 gene sequence and of a leucine zipper
motif in VP4 among five US bluetongue viruses.";
RL Virology 195:772-779(1993).
CC -!- FUNCTION: THE VP4 PROTEIN IS ONE OF THE FIVE PROTEINS (WITH VP1,
VP3, VP6 AND VP7) WHICH FORM THE INNER CAPSID OF THE VIRUS.
CC -!- SIMILARITY: BELONGS TO THE REOVIRUSES VP4 FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L08638; AAA42825.1; -.
CC Core protein.
KW SEQUENCE 644 AA; 75443 MW; 578503CFCD9B37D6 CRC64;
SQ

Query Match 8.1%; Score 84; DB 1; Length 644;
Best Local Similarity 22.8%; Pred. No. 2.8;
Matches 53; Conservative 16; Mismatches 73; Indels 90; Gaps 9;

QY 8 NDATVTPGLGEYEPPEPPNNTMRAENGRRPHHPYDAKQVS-----EYSCELHYTR 62
DB 141 NDANPRHLKHYGMPEPP---LYMEYAEIGTRFDDEPTDEKLVSLMDYIVYSAEHHVYV- 196
QY 63 FLTDGCRSAKPVTELVCSCGQC-GPARLL-----PNAIGRVKWRPNPGRDPFCIPDR 113
DB 197 -----GCCGLRTLMQFKRSGRFRRLVHWYDPIAPE--CSDPN 234
QY 114 YRAQRVOLLCPGAAPRSRKRLVASCCKELTRF-----HNOSELKDFGPETARPQKG 167
DB 235 VIVHNIWV-----DSKKILKMHNFKRVRLFTWVSDRSQMDHWEWETTRFAED 286
QY 168 RK-----PRPGAKANQAELEN 183
DB 287 RLGEIAYEMGGAFSSALIKHRIPNSKDEYHCISTYLFQPGADADMYELRN 338

RESULT 13
UL49_HSV11 STANDARD; PRT; 301 AA.
ID UL49_HSV11
AC P10233;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 01-DEC-1992 (Rel. 24, Last annotation update)
DE Tegument protein UL49.
GN UL49.
OS Herpes simplex virus (type 1 / strain 17).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88274327; PubMed=2839594;
RA McGeoch D.J., Dalrymple M.A., Davison A.J., Dolan A., Frame M.C.,

```

RA Clegg S.M., Cobley V.E., Cole C.G., Collier R.E., Connor R.,
RA Conroy P., Corby N.R., Coville G.J., Cox A.V., Davis J., Dawson E.,
RA Dhami P.D., Dockree C., Dodsorth S.J., Durbin R.M., Ellington A.G.,
RA Evans K.L., Fey J.M., Fleming K., French L., Garner A.A.,
RA Gilbert J.G.R., Goward M.E., Graffham D.V., Griffiths M.N.D., Hall C.,
RA Hall R.E., Hall-Tamlyn G., Heathcott R.W., Ho S., Holmes S.,
RA Hunt S.B., Jones M.C., Kerahaw J., Kimberley A.M., King A.,
RA Lait G.K., Langford C.F., Leversha M.A., Lloyd C., Lloyd D.M.,
RA Martyn I.D., Mashreghi-Mohammadi M., Matthews L.H., Mccann O.T.,
RA McElay J., McLaren S., McMurray A.A., Milne S.A., Mortimore B.J.,
RA Odell C.N., Pavitt R., Pearce A.V., Pearson D., Phillimore B.J.C.T.,
RA Phillips S.H., Plumb R.W., Ramsay H., Ramsey V., Rogers L., Ross M.T.,
RA Scott C.E., Sehra H.K., Skuce C.D., Smalley S., Smith M.L.,
RA Soderlund C., Spragon L., Steward C.A., Sulston J.E., Swann R.M.,
RA Vaudin M., Wall M., Wallis J.M., Whiteley M.N., Willey D.L.,
RA Williams L., Williams S.A., Williamson H., Wilmer T.E., Wilming L.,
RA Wright C.L., Hubbard T., Bentley D.R., Beck S., Rogers J., Shimizu N.,
RA Minoshima S., Kawasaki K., Sasaki T., Asakawa S., Kudoh J.,
RA Shintani A., Shibuya K., Yoshizaki Y., Aoki N., Mitsuyama S.,
RA Roe B.A., Chen F., Chu L., Crabtree J., Deschamps S., Do A., Do T.,
RA Dorman A., Fang F., Fu Y., Hu P., Hua A., Kenton S., Lai H., Lao H.I.,
RA Lewis J., Lewis S., Lin S.-P., Loh P., Mala J.E., Nguyen T., Pan H.,
RA Phan S., Qi S., Qian Y., Ray L., Ren Q., Shauli S., Sloan D., Song L.,
RA Wang Q., Wang Y., Wang Z., White J., Willingham D., Wu H., Yao Z.,
RA Zhan M., Zhang G., Chiesoe S., Murray J., Miller N., Minx P.,
RA Fulton R., Johnson D., Benis G., Bentley D., Bradshaw H., Bourne S.,
RA Cordes M., Du Z., Fulton L., Goela D., Graves T., Hawkins J.,
RA Hinds K., Kemp K., Latreille P., Layman D., Ozerky P., Rohlfing T.,
RA Scheet P., Walker C., Wamley A., Wohldmann P., Pepin K., Nelson J.,
RA Korf I., Bedell J.A., Hillier L., Mardis E., Waterston R., Wilson R.,
RA Emanuel B.S., Shaikh T., Kurauchi H., Saitta S., Budarf M.L.,
RA Mcdermid H.E., Johnson A., Wong A.C.C., Morrow B.E., Edelman L.,
RA Kim U.J., Shizuya H., Simon M.I., Dumanski J.P., Peyrard M., Kedra D.,
RA Serousi E., Franssen I., Tapia I., Bruder C.E., O'Brien K.P.,
RA Wilkinson P., Bodenteich A., Hartman K., Hu X., Khan A.S., Lane L.,
RA Tilahun Y., Wright H.,
RA "The DNA sequence of human chromosome 22."
RA Nature 402:489-495(1999).
RA [3]
RA SEQUENCE FROM N.A.
RA TISSUE=Muscle;
RA Strausberg R.;
RA Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RA -1- FUNCTION: May be involved in p53-induced cell cycle arrest in G2/M
RA phase by interfering with microtubule rearrangements that are
RA required to enter mitosis. Overexpression delays G2/M phase
RA progression.
RA -1- SUBCELLULAR LOCATION: Cytoplasmic. Associated with microtubules.
RA -1- DEVELOPMENTAL STAGE: Expressed in G2/M phase. Not detected in
RA quiescent cells.
RA -1- PTM: Phosphorylated in mitosis (By similarity).
RA -----
RA This SWISS-PROT entry is copyright. It is produced through a collaboration
RA between the Swiss Institute of Bioinformatics and the EMBL outstation -
RA the European Bioinformatics Institute. There are no restrictions on its
RA use by non-profit institutions as long as its content is in no way
RA modified and this statement is not removed. Usage by and for commercial
RA entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
RA or send an email to license@isb-sib.ch).
RA -----
RA EMBL; AF223408; AAF31459.1; -
RA EMBL; AL031588; CAB38415.1; -
RA EMBL; AL022325; CAB63079.1; -
RA EMBL; BC006325; AAH06325.1; -
RA KW Microtubules; Phosphorylation.
RA DOMAIN 22 27 POLY-SER.
RA CONFLICT 259 259 V -> I (IN REF. 1).
RA CONFLICT 506 506 R -> W (IN REF. 2; CAB38415).
RA SEQUENCE 720 AA; 76614 MW; ACD91CCCD008A89C CRC64;
RA -----
RA Query Match
RA Best Local Similarity 8.7%; Score 90; DB 1; Length 720;
RA Matches 34; Conservative 20; Mismatches 36; Indels 60; Gaps 6;

QY 68 PCESAPVTVLCVSCQCPARLLP-----NAIGRVKWMRPNPGDFRCIPDRYRAQRV----- 119
DB 341 PANSSRPLSNISKSGRMGMFAMLRPALPAGPVGASW-----QAARVDVSE 385
QY 120 ----QLLCPGGAAP-----RSRKVLVASC---KCRILT 146
DB 386 LAAEQLTAPPSASPTQPTPEGGGWLNSCAWESSQLNKTIRRRDSCLSNKTVMK 445
QY 147 RHNHSEKDF-----GPTARPKQGRKRP 172
DB 446 TPTNQFKPKFSGESPDSTPKLSRAQRP 475
RESULT 10
LSHB TRIVU
ID LSHB TRIVU STANDARD; PRT; 141 AA.
AC O46482;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Lutropin beta chain precursor (lutinizing hormone beta subunit) (LSH-
DE beta) (LSH-B) (LH-B).
GN LHB
OS Trichosurus vulpecula (Brush-tailed possum).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Metatheria; Diprotodontia; Phalangeridae; Trichosurus.
OX NCBI_TaxID=9337;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pituitary;
RA MEDLINE=98345424; PubMed=9680384;
RA Harrison G.A., Deane B.M., Cooper D.W.;
RA "cDNA cloning of lutinizing hormone subunits from brushtail possum
RA and red kangaroo".
RA Mamm. Genome 9:638-642(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA Lawrence S.B., McNatty K.P., Fidler A.E.;
RA Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
RA -1- FUNCTION: PROMOTES SPERMATOGENESIS AND OVULATION BY STIMULATING
RA THE TESTES AND OVARIES TO SYNTHESIZE STEROIDS.
RA -1- SUBUNIT: HETERODIMER OF A COMMON ALPHA CHAIN AND A UNIQUE BETA
RA CHAIN WHICH CONFERES BIOLOGICAL SPECIFICITY TO THYROTROPIN,
RA LUTROPIN, FOLLITROPIN AND GONADOTROPIN.
RA -1- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES BETA CHAIN
RA FAMILY.
RA -----
RA This SWISS-PROT entry is copyright. It is produced through a collaboration
RA between the Swiss Institute of Bioinformatics and the EMBL outstation -
RA the European Bioinformatics Institute. There are no restrictions on its
RA use by non-profit institutions as long as its content is in no way
RA modified and this statement is not removed. Usage by and for commercial
RA entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
RA or send an email to license@isb-sib.ch).
RA -----
RA EMBL; AF017448; AAC96019.1; -
RA EMBL; AF090388; AAC63526.1; -
RA HSSP; P01233; 1XUL.
RA InterPro; IPR000359; Cys_knot.
RA InterPro; IPR001545; Gly_hormone.
RA Pfam; PF00007; Cys_knot; 1.
RA SMART; SM00068; GHB; 1.
RA PROSITE; PS00261; GLYCO_HORMONE_BETA_1; 1.
RA PROSITE; PS00689; GLYCO_HORMONE_BETA_2; 1.
RA Hormone; Signal; Glycoprotein.
RA SIGNAL 1 22
RA CHAIN 23 141 LUTROPIN BETA CHAIN.
RA DISULFID 30 73 BY SIMILARITY.
RA DISULFID 44 93 BY SIMILARITY.
RA DISULFID 47 131 BY SIMILARITY.
RA DISULFID 55 109 BY SIMILARITY.
RA DISULFID 59 111 BY SIMILARITY.

RA Beighton P., Mulligan J.T.,
 RT "Bone dysplasia sclerosteosis results from loss of the SOST gene
 RL product, a novel cysteine knot-containing protein.",
 CC Am. J. Hum. Genet. 68:577-589(2001).
 CC -!- FUNCTION: Seems to play a role in bone homeostasis (By
 CC similarity).
 CC -!- SUBCELLULAR LOCATION: Secreted (Potential).
 CC -!- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTINE KNOT-LIKE (CTCK) DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AF326738; AAK13453.1; -
 DR InterPro: IPR000359; Cys_knot.
 DR SMART: SM00041; CT; 1.
 DR PROSITE: PS01185; CTCK_1; FALSE_NEG.
 DR PROSITE: PS01225; CTCK_2; FALSE_NEG.
 KW Glycoprotein.
 FT NON TER 1 1
 FT DOMAIN 51 141 CTCK.
 FT CARBOHYD 22 22 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 144 144 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT NON TER 176 176
 SQ SEQUENCE 176 AA; 19743 MW; 35F0CA61A25F4DB CRC64;

Query Match 81.3%; Score 839; DB 1; Length 176;
 Best Local Similarity 86.4%; Pred. No. 7e-69;
 Matches 152; Conservative 11; Mismatches 9; Indels 4; Gaps 2;
 QY 8 NDATEVIFGLGEYEPPEPE-NNQTMRAENGRRPPHPYDAKDVSEVSCRELHYTRFLTD 66
 DB 1 NDATEIIPGLGEYEPPEPELNKNKTMRAENGRRPPHPYDAKDVSEVSCRELHYTRFLTD 60
 QY 67 GPCHSAKPVTELVCSGCGGPARLLPNAIGRWKWRPNPGDFRCIPDPRYRAQRVQLCPGG 126
 DB 61 GPCHSAKPVTELVCSGCGGPARLLPNAIGRWKWRPNPGDFRCIPDPRYRAQRVQLCPGG 120
 QY 127 AAPRSKRVLVASCKKRLTRFNQSELKDFGPETARPQKGRKRP---GAKANQA 179
 DB 121 AAPARKRVLVASCKKRLTRFNQSELKDFGPETARPQKGRKRP---GAKANQA 176

RESULT 6
 LSHB_PHYCA STANDARD; PRT; 118 AA.
 AC P25330;
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 01-MAY-1992 (Rel. 22, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Lutropin beta chain (lutinizizing hormone beta subunit) (LSH-beta)
 DE (LSH-B) (LH-B).
 GN LHB.
 OS Physeter catodon (Sperm whale) (Physeter macrocephalus).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Cetacea; Odontoceti;
 OC Physeteridae; Physeter.
 OX NCBI_TaxID=9755;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=87032654; PubMed=3771098;
 RA Pankov Y.A., Karasyov V.S.;
 RT "Primary structure of sperm whale lutinizizing hormone.";
 RL Int. J. Pept. Protein Res. 28:124-129(1986).
 RN [2]
 RP SEQUENCE.
 RX MEDLINE=84281133; PubMed=6466737;
 RA Pankov Y.A., Karasev V.S.;
 RT "Lutinizizing hormone of the sperm whale. Amino acid sequences of

RT reduced and carboxymethylated beta-subunits.",
 RL Biochimia 49:1004-1018(1984).
 CC -!- FUNCTION: PROMOTES SPERMATOGENESIS AND OVULATION BY STIMULATING
 CC THE TESTES AND OVARIES TO SYNTHESIZE STEROIDS.
 CC -!- SUBUNIT: HETERODIMER OF A COMMON ALPHA CHAIN AND A UNIQUE BETA
 CC CHAIN WHICH CONFERS BIOLOGICAL SPECIFICITY TO THYROTROPIN,
 CC LUTROPIN, FOLLITROPIN AND GONADOTROPIN.
 CC -!- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES BETA CHAIN
 CC FAMILY.
 DR PIR: P0141; P0141.
 DR HSP; P01233; 1X0L.
 DR InterPro: IPR000359; Cys_knot.
 DR InterPro: IPR002400; GF_cysknot.
 DR InterPro: IPR001545; Gly_hormoneB.
 DR Pfam: PF00007; Cys_knot_1.
 DR PRINTS: PR00438; GFCYSKNOT.
 DR SMART: SM00068; GHB; 1.
 DR PROSITE: PS00261; GLYCO_HORMONE_BETA_1; 1.
 DR PROSITE: PS00689; GLYCO_HORMONE_BETA_2; 1.
 KW Hormone; Glycoprotein.
 FT DISULFID 9 37
 FT DISULFID 23 72 BY SIMILARITY.
 FT DISULFID 26 110 BY SIMILARITY.
 FT DISULFID 34 38 BY SIMILARITY.
 FT DISULFID 38 90 BY SIMILARITY.
 FT DISULFID 93 100 BY SIMILARITY.
 FT CARBOHYD 13 13 N-LINKED (GLCNAC. . .).
 SQ SEQUENCE 118 AA; 12412 MW; 81177A56382F15E7 CRC64;
 Query Match 9.0%; Score 93; DB 1; Length 118;
 Best Local Similarity 30.1%; Pred. No. 0.079;
 Matches 37; Conservative 13; Mismatches 55; Indels 18; Gaps 6;
 QY 55 CRELHYTRFLTDGPRCHSAKPVTELVCSGCG-PPARLLPNAIGRWKWRPNPGDFRCIPDR 113
 DB 9 CRPINATLAAQNZACFCVITFTTISICAGYCPSVVRVLPAL-----PPVZPVCTYRQ 61
 QY 114 YRAQRVQLL-CPGAPR-PR-SRKVLVASCCKRLTRFNQSELKDFGPETARPQK-GRKP 170
 DB 62 LRFASIRLPCGPPGVNPMVSFPVALSCHGCPCLSS-----SDCGPGRAQPLACNRSP 114
 QY 171 RPG 173
 DB 115 RPG 117

RESULT 7
 LSHB_MELGA STANDARD; PRT; 159 AA.
 AC P45646;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Lutropin beta chain precursor (Luteinizing hormone beta subunit) (LSH-
 DE beta) (LSH-B) (LH-B).
 GN LHB.
 OS Meleagris gallopavo (Common turkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Meleagrididae; Meleagris.
 OX NCBI_TaxID=9103;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Pituitary;
 RX MEDLINE=95290073; PubMed=7772235;
 RA You S., Foster L.K., Silsby J.L., el Halawani M.E., Foster D.N.;
 RT "Sequence analysis of the turkey LH beta subunit and its regulation
 RT by gonadotrophin-releasing hormone and prolactin in cultured
 RT pituitary cells.";
 RL J. Mol. Endocrinol. 14:117-129(1995).
 CC -!- FUNCTION: PROMOTES SPERMATOGENESIS AND OVULATION BY STIMULATING
 CC THE TESTES AND OVARIES TO SYNTHESIZE STEROIDS.
 CC -!- SUBUNIT: HETERODIMER OF A COMMON ALPHA CHAIN AND A UNIQUE BETA
 CC CHAIN WHICH CONFERS BIOLOGICAL SPECIFICITY TO THYROTROPIN,

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

CC ENBL; AF326742; AAK13457.1; -.

CC InterPro; IPR000359; Cys_knot.

CC SMART; SM00041; CT; 1.

CC PROSITE; PS01185; CTCK_1; FALSE_NEG.

CC PROSITE; PS01225; CTCK_2; FALSE_NEG.

CC KW Signal; Glycoprotein 23 POTENTIAL.

FT CHAIN 24 213 SCLEROSTIN.

FT DOMAIN 82 172 CTCK.

FT CARBOHYD 53 53 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 175 175 N-LINKED (GLCNAC. . .) (POTENTIAL).

CC SEQUENCE 213 AA; 23908 MW; 6DA7B5EDE674728A CRC64;

Query Match 92.2%; Score 951.5; DB 1; Length 213;

Best Local Similarity 90.0%; Pred. No. 6.1e-79; Indels 5; Gaps 2;

Matches 171; Conservative 10; Mismatches 4;

Qy 1 QGWAFRNDATVPIGLGEYPEPPP--ENNQTWNRAENGGRPHHPYDADKDVSEYSCREL 58

Db 24 QGWAFRNDATVPIGLGEYPEPPP--ENNQTWNRAENGGRPHHPYDADKDVSEYSCREL 83

Qy 59 HYTRFLTDGPCRSAPVTVLCVSCGCGPARLLPNAIGRVKWRPENGDFRCIPDRYRAQR 118

Db 84 HPTRYVTDGPCRSAPVTVLCVSCGCGPARLLPNAIGRVKWRPENGDFRCIPDRYRAQR 143

Qy 119 VOLLCPGGAAPRSKRVLRVASCCKRLTRFNQSELKDFGPETARPOKGRKPRP---GAK 175

Db 144 VOLLCPGGAAPRSKRVLRVASCCKRLTRFNQSELKDFGPETARPOKGRKPRP---GAK 203

Qy 176 ANQAELENAY 185

Db 204 ANQAELENAY 213

Query Match 90.4%; Score 932.5; DB 1; Length 213;

Best Local Similarity 88.4%; Pred. No. 3.2e-77; Indels 5; Gaps 2;

Matches 168; Conservative 10; Mismatches 7;

Qy 1 QGWAFRNDATVPIGLGEYPEPPP--ENNQTWNRAENGGRPHHPYDADKDVSEYSCREL 58

Db 24 QGWAFRNDATVPIGLGEYPEPPP--ENNQTWNRAENGGRPHHPYDADKDVSEYSCREL 83

Qy 59 HYTRFLTDGPCRSAPVTVLCVSCGCGPARLLPNAIGRVKWRPENGDFRCIPDRYRAQR 118

Db 84 HPTRYVTDGPCRSAPVTVLCVSCGCGPARLLPNAIGRVKWRPENGDFRCIPDRYRAQR 143

Qy 119 VOLLCPGGAAPRSKRVLRVASCCKRLTRFNQSELKDFGPETARPOKGRKPRP---GAK 175

Db 144 VOLLCPGGAAPRSKRVLRVASCCKRLTRFNQSELKDFGPETARPOKGRKPRP---GAK 203

Qy 176 ANQAELENAY 185

Db 204 ANQAELENAY 213

RESULT 5

SOST BOVIN STANDARD; PRT; 176 AA.

ID SOST BOVIN STANDARD; PRT; 176 AA.

AC Q9BQ79;

DT 15-JUN-2002 (Rel. 41, Created)

DT 15-JUN-2002 (Rel. 41, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Sclerostin (Fragment).

GN SOST.

OS Bos taurus (Bovine).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

OC Bovidae; Bovinae; Bos.

OX NCBI_TaxID=9913;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=21090529; PubMed=11179006;

RA Brunkow M.E., Gardner J.C., Van Ness J., Paepers B.W., Kovacevich B.R.,

RA Gillett L., Colbert T., Tacconi P., Galas D., Hamersma H.,

RA Beighton P., Mulligan J.T.;

RT "Bone dysplasia sclerostosis results from loss of the SOST gene product, a novel cysteine knot-containing protein.";

RL Am. J. Hum. Genet. 68:577-589(2001).

CC -!- FUNCTION: Seems to play a role in bone homeostasis.

CC -!- SUBCELLULAR LOCATION: Secreted (Potential).

CC -!- TISSUE SPECIFICITY: Widely expressed at low levels with highest

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; AF326740; AAK13455.1; -
 CC EMBL; AK017295; BAB30678.1; -
 CC EMBL; AF326737; AAK13452.1; -
 CC MGD; MGI:1921749; SOST.
 CC InterPro; IPR000359; Cys_knot.
 CC PROSITE; PS01185; CTCK_1; FALSE_NEG.
 CC PROSITE; PS01225; CTCK_2; FALSE_NEG.
 CC Signal; Glycoprotein.
 CC SIGNAL 1 23 POTENTIAL.
 CC CHAIN 24 211 SCLEROSTIN.
 CC DOMAIN 80 170 CTCK.
 CC CARBOHYD 51 51 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CARBOHYD 173 173 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CONFLICT 72 72 G -> D (IN REF. 2).
 CC SEQUENCE 211 AA; 23443 MW; AEB094E358E34961 CRC64;

Query Match 98.2%; Score 1013.5; DB 1; Length 211;
 Best Local Similarity 97.9%; Pred. No. 1.6e-84;
 Matches 184; Conservative 0; Mismatches 1; Indels 3; Gaps 1;

Qy 1 QGWAFRNDATVPIGLGEYPPPPENNQTMNRAENGRRPHHPYDAKDVSYSRELHY 60
 Db 24 QGWAFRNDATVPIGLGEYPPPPENNQTMNRAENGRRPHHPYDAKDVSYSRELHY 83
 Qy 61 TRFLTDGPCRSAPVTVELVCSGCGPARLLPNAIGRVKWRPNPDPFCIPDRYRAQRVQ 120
 Db 84 TRFLTDGPCRSAPVTVELVCSGCGPARLLPNAIGRVKWRPNPDPFCIPDRYRAQRVQ 143
 Qy 121 LCPGGAAPRSKRVLRVASCCKRLTRFNQSELKDFGFPETARPQGRKPRP---GAKAN 177
 Db 144 LCPGGAAPRSKRVLRVASCCKRLTRFNQSELKDFGFPETARPQGRKPRP---GAKAN 203
 Qy 178 QAELENAY 185
 Db 204 QAELENAY 211

RESULT 2
 SOST_RAT
 ID SOST_RAT STANDARD; PRT; 213 AA.
 AC Q99567;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DE Sclerostin precursor.
 GN SOST.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley;
 RX MEDLINE=21090529; PubMed=11179006;
 RA Brunkow M.E., Gardner J.C., Van Ness J., Paepers B.W., Kovacevich B.R.,
 RA Proll S., Skonier J.E., Zhao L., Sabo P.J., Fu Y.H., Alisch R.S.,
 RA Gillett L., Colbert T., Tacconi P., Galas D., Hamersma H.,
 RA Beighton P., Mulligan J.T.;
 RT "Bone dysplasia sclerosteosis results from loss of the SOST gene
 RT product, a novel cysteine knot-containing protein.";
 RL Am. J. Hum. Genet. 68:577-589(2001).
 CC -!- FUNCTION: Seems to play a role in bone homeostasis (By
 CC similarity).
 CC -!- SUBCELLULAR LOCATION: Secreted (Potential).
 CC -!- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTINE KNOT-LIKE (CTCK) DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; AF326741; AAK13456.1; -
 CC InterPro; IPR000359; Cys_knot.
 CC PROSITE; PS01185; CTCK_1; FALSE_NEG.
 CC PROSITE; PS01225; CTCK_2; FALSE_NEG.
 CC Signal; Glycoprotein.
 CC SIGNAL 1 28 POTENTIAL.
 CC CHAIN 29 213 SCLEROSTIN.
 CC DOMAIN 82 172 CTCK.
 CC CARBOHYD 53 53 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CARBOHYD 175 175 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC SEQUENCE 213 AA; 23974 MW; 6C56C878BCDC684B CRC64;

Query Match 95.2%; Score 982.5; DB 1; Length 213;
 Best Local Similarity 94.2%; Pred. No. 9.9e-82;
 Matches 179; Conservative 3; Mismatches 3; Indels 5; Gaps 2;

Qy 1 QGWAFRNDATVPIGLGEYPPPPENNQTMNRAENGRRPHHPYDAKDVSYSREL 58
 Db 24 QGWAFRNDATVPIGLGEYPPPPENNQTMNRAENGRRPHHPYDAKDVSYSREL 83
 Qy 59 HYTRFLTDGPCRSAPVTVELVCSGCGPARLLPNAIGRVKWRPNPDPFCIPDRYRAQR 118
 Db 84 HYTRFLTDGPCRSAPVTVELVCSGCGPARLLPNAIGRVKWRPNPDPFCIPDRYRAQR 143
 Qy 119 VOLLCPGGAAPRSKRVLRVASCCKRLTRFNQSELKDFGFPETARPQGRKPRP---GAK 175
 Db 144 VOLLCPGGAAPRSKRVLRVASCCKRLTRFNQSELKDFGFPETARPQGRKPRP---GAK 203
 Qy 176 ANQAELENAY 185
 Db 204 ANQAELENAY 213

RESULT 3
 SOST_CERAE
 ID SOST_CERAE STANDARD; PRT; 213 AA.
 AC Q9BG78;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DE Sclerostin precursor.
 GN SOST.
 OS Cercopithecus aethiops (Green monkey) (Grivet).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecinae; Cercopithecus.
 OX NCBI_TaxID=9534;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21090529; PubMed=11179006;

RA Brunkow M.E., Gardner J.C., Van Ness J., Paepers B.W., Kovacevich B.R.,
 RA Proll S., Skonier J.E., Zhao L., Sabo P.J., Fu Y.H., Alisch R.S.,
 RA Gillett L., Colbert T., Tacconi P., Galas D., Hamersma H.,
 RA Beighton P., Mulligan J.T.;
 RT "Bone dysplasia sclerosteosis results from loss of the SOST gene
 RT product, a novel cysteine knot-containing protein.";
 RL Am. J. Hum. Genet. 68:577-589(2001).
 CC -!- FUNCTION: Seems to play a role in bone homeostasis (By
 CC similarity).
 CC -!- SUBCELLULAR LOCATION: Secreted (Potential).
 CC -!- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTINE KNOT-LIKE (CTCK) DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

GenCore version 5.1.4.p5 4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 28, 2003, 14:11:46 ; Search time 14.3067 Seconds
(without alignments)

536.332 Million cell updates/sec

Title: US-09-867-274-4

Perfect score: 1032

Sequence: 1 CQWQAFNRDATEVPGIGY.....KGRKPRPGAKAQAELNAY 185

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1013.5	98.2	211	1 SOST_MOUSE	Q9968 mus musculus
2	982.5	95.2	213	1 SOST_RAT	Q9967 rattus norv
3	951.5	92.2	213	1 SOST_CERAE	Q9b78 cercopithe
4	932.5	90.4	213	1 SOST_HUMAN	Q9b74 homo sapien
5	839	81.3	176	1 SOST_BOVIN	Q9b79 bos taurus
6	93	9.0	118	1 LSHB_PHYCA	P25330 physeter ca
7	92.5	9.0	159	1 LSHB_MELGA	P45646 meleagris g
8	91	8.8	118	1 LSHB_BALAC	P33088 balaenopter
9	90	8.7	720	1 GTSE_HUMAN	Q9ny23 homo sapien
10	86.5	8.4	141	1 LSHB_TRIVU	O46482 trichosurus
11	84	8.1	644	1 VP4_BTIV1	P33428 bluetongue
12	84	8.1	644	1 VP4_BTIV13	P33429 bluetongue
13	82	7.9	301	1 U49_HSV11	P10233 herpes simp
14	82	7.9	2182	1 CAB1_RAT	O88480 rattus norv
15	80	7.8	141	1 LSHB_CERSI	O77835 ceratotheri
16	80	7.8	654	1 VP4_BTIV10	P07132 bluetongue
17	79.5	7.7	165	1 LSHB_CONJA	P45657 coturnix co
18	78.5	7.6	625	1 DUS8_HUMAN	O13202 homo sapien
19	78.5	7.6	1942	1 Y054_HUMAN	P42694 homo sapien
20	77	7.5	470	1 KL4F_HUMAN	O43474 homo sapien
21	76.5	7.4	138	1 LSHB_MACRU	O46483 macropus ru
22	76.5	7.4	165	1 CSHB_HUMAN	P01233 homo sapien
23	76	7.4	143	1 LSHB_FELCA	O77805 felis silve
24	76	7.4	180	1 DAN_HUMAN	P41271 homo sapien
25	76	7.4	232	1 HXB5_XENLA	P31272 xenopus lae
26	76	7.4	888	1 LSHB_MOUSE	P08923 mus musculus
27	74.5	7.2	128	1 LSHB_STRCA	P80664 struthio ca
28	74	7.2	178	1 DAN_RAT	Q06880 rattus norv
29	74	7.2	379	1 CVR6_MOUSE	P18406 mus musculus
30	74	7.2	381	1 CVR6_HUMAN	O06622 homo sapien
31	74	7.2	930	1 SM6C_HUMAN	Q9h3t2 homo sapien
32	74	7.2	1229	1 N121_HUMAN	Q9y2n3 homo sapien
33	73.5	7.1	623	1 EXAA_PSEAE	Q9z4j7 pseudomonas

ALIGNMENTS

RESULT 1
SOST_MOUSE
ID SOST_MOUSE STANDARD; PRT; 211 AA.
AC Q99P68; Q9D3L7;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Sclerostin precursor.
GN SOST.

OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]

SEQUENCE FROM N.A.

RC STRAIN=129/SVJ;
RX MEDLINE=21090529; PubMed=11179006;
RA Brunkow M.E., Gardner J.C., Van Ness J., Paepker B.W., Kovacevich B.R.,
RA Proll S., Skonier J.E., Zhao L., Sabo P.J., Fu Y.H., Alisch R.S.,
RA Gillett L., Colbert I., Tacconi P., Galas D., Hamersma H.,

RA Beighton P., Mulligan J.T.;
RA "bone dysplasia sclerosteosis results from loss of the SOST gene product, a novel cystine knot-containing protein.",
RL Am. J. Hum. Genet. 68:577-589(2001).
RN [2]

SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Schriml L.M., Stauble F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Kamiya M., Lee N.H.,
RA Gustingich S., Hall D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Sakamoto N.,
RA Nordone P., Ringwald M., Rodriguez I., Sakamoto N., Storch K.-F.,
RA Sasaki H., Sato K., Schoenbach C., Seva T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;

RT "Functional annotation of a full-length mouse cDNA collection";
RL Nature 409:685-690(2001).

CC -!- FUNCTION: Seems to play a role in bone homeostasis (By similarity).

CC -!- SUBCELLULAR LOCATION: Secreted (Potential).

CC -!- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTINE KNOT-LIKE (CTCK) DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its

A;Cross-references: EMBL:AF061947; NID:g3232606; PID:g3232607; PIDN:AAC40176.1

Query Match 7.9%; Score 82; DB 2; Length 2182;
Best Local Similarity 22.8%; Pred. No. 43;
Matches 50; Conservative 17; Mismatches 68; Indels 84; Gaps 9;

QY 4 QAFRDATEVIP-----GLGEYPEPP-----PENNQWRAENG-----R 39
DB 1693 KASPEDQESLPHKPLPLADSGPGPEGKVGFLHQLPVATDRNTDQGGEPKDKER 1752
QY 40 PPHHPVDKDVSEYSCR-ELHYT-RFLTDGCRSAKP-----VTELVCSGQCGPAR 88
DB 1753 PVGTEPMGTGETAARHPDLEPTPLLPGRPPDRDGPESRASAELSLELSISTRQCPAP 1812
QY 89 LLPNAIGRVKWWRNPGDPFCIPDKRYAQRVOLLCPGGAAPRSKRVLVASCKKRLTRF 148
DB 1813 LVPSPV-----TPTTAAPTMTGARAAGH-----1835
QY 149 HNOSELKDFGPETA--RPQGRKPRPOKAKANOAELENAY 185
DB 1836 -----PEEAPPRPNRKRKLQDTEGSKTLLDAY 1864

RESULT 11

S71793 head-inducing factor Cerberus - African clawed frog

C;Species: Xenopus laevis (African clawed frog)

C;Date: 12-Feb-1998 #sequence_revision 13-Mar-1998 #text_change 21-Jul-2000

C;Accession: S71793

R;Boumeester, T.; Kim, S.H.; Sasaki, Y.; Lu, B.; de Robertis, E.M.

A;Title: Cerberus is a head-inducing secreted factor expressed in the anterior endoderm

A;Reference number: S71793; MUID:96338220; PMID:8757128

A;Accession: S71793

A;Status: preliminary; nucleic acid sequence not shown

A;Molecule type: mRNA

A;Residues: 1-270 <BOU>

A;Cross-references: EMBL:U64831; NID:g1513087; PIDN:AAC60012.1; PID:g1513088

Query Match

Best Local Similarity 7.8%; Score 80.5; DB 2; Length 270;

Matches 28; Conservative 27; Mismatches 53; Indels 25; Gaps 7;

QY 38 GRPPH--HPYDAKDVSRELYHTRFLTDGCRSAKPVTCLVCSGCGCPARLLPNAIG 95
DB 150 GAPQNTSHGSKAQEIMKEACKLTPTQNLVHNC--DRMVIONNLCFGKC-----197

QY 96 RVKWWPNPGDPR-----CIPDRYRAQRVOLLCPGGAAPRSKRVLVASCKK-RLTRFH 149

DB 198 -ISLHVPNOODRRNTCSHCLPSKFTLNHLNCT-GSKNVVKKVMVVEECTCEAHKSNFH 255

QY 150 NQSELKDFGPETA 162

DB 256 QTAQ----FNWDT 265

RESULT 12

P4XRBV

minor inner core protein VP4 - bluetongue virus (serotype 10, American isolate)

C;Species: bluetongue virus

C;Date: 30-Sep-1989 #sequence_revision 30-Sep-1989 #text_change 10-Nov-1995

C;Accession: A26862; S10537

R;Yu, Y.; Fukusho, A.; Roy, P.

Nucleic Acids Res. 15, 7206, 1987

A;Title: Nucleotide sequence of the VP4 core protein gene (M4 RNA) of US bluetongue virus

A;Reference number: A26862; MUID:88015582; PMID:2921504

A;Accession: A26862

A;Molecule type: genomic RNA

A;Residues: 1-654 <YUY>

R;Roy, P.; Marshall, J.J.A.; French, T.J.

Curr. Top. Microbiol. Immunol. 162, 43-87, 1990

A;Title: Structure of the bluetongue virus genome and its encoded proteins.

A;Reference number: S10534; MUID:90345726; PMID:2166648

A;Accession: S10537

A;Status: preliminary

A;Molecule type: genomic RNA

A;Residues: 1-654 <ROY>

C;Genetics:

A;Map position: segment 4

C;Superfamily: bluetongue virus core protein VP4

C;Keywords: core protein

Query Match

Best Local Similarity 7.8%; Score 80; DB 1; Length 654;

Matches 53; Conservative 16; Mismatches 73; Indels 90; Gaps 9;

QY 8 NDATVIEPGLGEYPEPPENNTWRAENGGRPPHHYDAKDV-----EYSCRELHYTR 62

DB 141 NDANRELHKIYGMPIPP---LYMEYAEIGTRFDDEPTDEKLVSMHYIYVSAEYHYI- 196

QY 63 FLTDGCRSAKPVTCLVCSGCGCPARLL-----PNAIGRVKW--WRPNPGDPRFCIPDR 113

DB 197 -----CGDLRLTLMQFKKSPGRFRVRLWHYVDPIAPB--CSDPN 234

QY 114 YRAQRVOLLCPGGAAPRSKRVLVASCKKRLTR-----FHNQSELKDFGPETARPQKG 167

DB 235 VIVHNIW-----DSKDKILKMNFLKRVPRPFIWDVSSDRSQNDHEWETTRFAD 286

QY 168 RK-----PRPGAKANOAELEN 183

DB 287 RLGEIAYEMGGAFSSALIKHRIPNSKDEYHCISTYLVFPQGDADADNYELRN 338

RESULT 13

AS7534

mucin SAC (clone L31) - human (fragment)

C;Species: Homo sapiens (man)

C;Date: 08-Feb-1996 #sequence_revision 08-Feb-1996 #text_change 20-Apr-2000

C;Accession: A57534

R;Lesuffleur, T.; Roche, F.; Hill, A.S.; Lacasa, M.; Fox, M.; Swallow, D.M.; Zweibaum, A.

J. Biol. Chem. 270, 13665-13673, 1995

A;Title: Characterization of a mucin cDNA clone isolated from HT-29 mucus-secreting cells

A;Reference number: A57534; MUID:95293957; PMID:7775418

A;Accession: A57534

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-1042 <LES>

A;Cross-references: GB:246314; NID:g1052607; PIDN:CAA68307.1; PID:g1052608

C;Genetics:

A;Gene: GDB:MUCSAC

A;Cross-references: GDB:454136; OMIM:158373

A;Map position: l1p15.5-11p15.5

C;Superfamily: von Willebrand factor type C repeat homology

F;678-746/Domain: von Willebrand factor type C repeat homology <WVC>

Query Match

Best Local Similarity 7.8%; Score 80; DB 2; Length 1042;

Matches 35; Conservative 16; Mismatches 55; Indels 58; Gaps 6;

QY 23 PPEPNNQTNWRAENGGRPPHHYDAKDVSEYSCRELHYTRFLTDGCRSAKPVTCLVCSG 82

DB 819 PPPYQNS-----TCAYYHRSLLIQOQCGSSSEPVRLAYCRG 855

QY 83 QCGPARLLPNAIGRVKWWRNPGDPR---CIPDRYRAQRVOLLCPGGAAPRSKRVLVAS 139

DB 856 NCGDSSSMYSLEG-----NTVEHRCQCCQLTSLRNVLHCTDGS--RAFSYTEVEE 907

QY 140 CKCKRLTRFHNQSELKDFGPETARPQGRK--PRPGAKANOAELE 182

DB 908 CGC-----MGRCPAPGDTQHSEEA 928

RESULT 14

I51242

luteinizing hormone beta-subunit - quail

C;Species: Coturnix coturnix (quail)

Mon Mar 31 09:44:36 2003

us-09-867-274-4.xpr

```

A;Introns: 294/2
A;Molecule type: DNA
A;Residues: 1-334 <BEV>
A;Cross-references: EMBL:AL021635
A;Experimental source: cultivar Columbia; BAC clone T12H17
C;Genetics:
A;Map position: 4
A;Introns: 139/1; 161/3; 205/3
A;Note: T12H17.160

Query Match      8.0%; Score 82.5; DB 2; Length 334;
Best Local Similarity 25.9%; Pred. No. 6;
Matches 28; Conservative 13; Mismatches 30; Indels 37; Gaps 4;

QY 21 PEPPENNQTMMRAENG-----GRPPHPYDAKDVSSEYSCRELHYTRFLTDGCR 70
Db 49 PPPPQNSFTPSAAMDGFSSGPIKRRGRPRKYGHGAAVT-----LSPNPIS 96

QY 71 SAKPVTELVC-----GQCGPARLLPNAIGRVKMW-----WRPN 103
Db 97 SAAPTSHVIDFSTSEKRGKMKPATPTPSSFIRPKYQVENLGWSPS 144

RESULT 9
WMBEF9
UL49 protein - human herpesvirus 1 (strain 17)
C;Species: human herpesvirus 1
C;Date: 31-Dec-1989 #sequence_revision 31-Dec-1989 #text_change 16-Jun-2000
C;Accession: D30089
R;McGeoch, D.J.; Dalrymple, M.A.; Davison, A.J.; Dolan, A.; Frame, M.C.; McNab, D.; Perry, J.; Gen. Virol. 69, 1531-1574, 1988
A;Title: The complete DNA sequence of the long unique region in the genome of herpes simi
A;Reference number: A30083; MUID:88274327; PMID:2839594
A;Accession: D30089
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-301 <MCG>
A;Cross-references: GB:X14112; NID:G1944536; PIDN:CAA32299.1; PID:G59549; GB:D00317
C;Genetics:
A;Gene: UL49
C;Superfamily: varicella-zoster virus gene 9 protein

Query Match      7.9%; Score 82; DB 1; Length 301;
Best Local Similarity 26.1%; Pred. No. 6;
Matches 43; Conservative 17; Mismatches 67; Indels 38; Gaps 10;

QY 21 PEPPENNQ---TMMRAENGRRPPHPYDAKDVSSEY-----SCRELHYTRFLTDGCRSAK 73
Db 33 PDSPDTSRRGALQTKRSRQGEVRFVQYDESDYALYGSSSEDEH-----PEVP-RTRR 86

QY 74 PVTELYCSGCGPARLLP-----NAIGRVKWRPNPGDFRCIPDRYRAQVQLLCPGAA 128
Db 87 PVSGAVLSGP-GEARAPPPAGSGGAGRTPTTAPRAP-----RTQRVATKAP--AA 134

QY 129 PRSRKRVLVASCKKCLTRFHQNSKLDKFGPETARPKGKPRPG 173
Db 135 PAETTR-----GRKSAQPESALPD-APASTATPRTSKTAPQG 171

RESULT 10
T14320
calcineurin inhibitor cain - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C;Accession: T14320
R;Lai, M.M.; Burnett, P.E.; Wolosker, H.; Blackshaw, S.; Snyder, S.H.
J. Biol. Chem. 273, 18325-18331, 1998
A;Title: Cain, a novel physiologic protein inhibitor of calcineurin.
A;Reference number: Z17979; MUID:98325042; PMID:9660798
A;Accession: T14320
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-2182 <LAI>

```


GenCore version 5.1.4_p5 4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 28, 2003, 14:13:46 ; Search time 26.64 Seconds
(without alignments)
667.600 Million cell updates/sec

Title: US-09-867-274-4

Perfect score: 1032

Sequence: 1 CQWQAFNRDATEVIFGLGEY.....KGRKPRPGAKANQALENAY 185

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_73: *
1: piri: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	249	24.1	134	2 T08710	hypothetical prote
2	93	9.0	118	2 PNO141	lutropin beta chai
3	92.5	9.0	159	2 I51373	lutropin beta chain
4	91	8.8	118	2 PNO139	lutropin beta chain
5	89.5	8.7	158	2 A61091	ribosomal protein
6	88	8.5	332	2 S46947	adenylate cyclase
7	84.5	8.2	1223	2 S29717	hypothetical prote
8	82.5	8.0	334	2 T04572	UL49 protein - hum
9	82	7.9	301	1 WMBEP9	calcineurin inhibi
10	82	7.9	2182	2 T14320	head-inducing fact
11	80.5	7.8	270	2 S71793	minor inner core p
12	80	7.8	654	1 P4XRBV	mucin SAC (clone L
13	80	7.8	1042	2 A57534	lutropin beta chain
14	79.5	7.7	166	2 I51242	hypothetical prote
15	79.5	7.7	601	2 T22025	protein P40E10.4 l
16	79.5	7.7	601	2 D89711	hypothetical prote
17	78.5	7.6	1142	2 T30272	protein CTRP - mal
18	78.5	7.6	2098	2 T18397	vascular endotheli
19	77.5	7.5	188	2 JC4680	ribosomal protein
20	77.5	7.5	349	2 S49606	hypothetical prote
21	77	7.5	834	2 T42702	choriognadotropin
22	76.5	7.4	165	1 KTHUB	neural induction h
23	76	7.4	232	2 A60083	cell adhesion mole
24	76	7.4	405	2 S33601	tyrosine kinase -
25	76	7.4	888	2 I58378	slit protein kinase
26	76	7.4	1469	2 B36665	mucin MUC5B, trach
27	75.5	7.3	3570	2 T45025	thioredoxin trxa l
28	75	7.3	322	2 AH3011	probable thioredox
29	75	7.3	331	2 G98272	

30 75 7.3 1736 2 T00391
31 74.5 7.2 128 2 S74085
32 74.5 7.2 275 2 T50578
33 74 7.2 178 2 I51824
34 74 7.2 178 2 A47291
35 74 7.2 269 2 D75631
36 74 7.2 294 2 T34537
37 74 7.2 379 2 A35669
38 74 7.2 866 2 T29197
39 73.5 7.1 145 2 I37231
40 73.5 7.1 402 2 T48729
41 73.5 7.1 623 2 B83399
42 73.5 7.1 650 2 A34498
43 73.5 7.1 701 2 T52384
44 73.5 7.1 825 2 JC4163
45 73 7.1 433 2 D84335

ALIGNMENTS

RESULT 1

T08710
hypothetical protein DKFZp564D206.1 - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 13-Aug-1999

C:Accession: T08710

R:Wambutt, R.; Heubner, D.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, March 1999

A:Reference number: 216471

A:Accession: T08710

A:Molecule type: mRNA

A:Residues: 1-134 <WAM>

A:Cross-references: EMBL:AL050024

A:Experimental source: fetal brain; clone DKFZp564D206

C:Genetics:

A:Note: DKFZp564D206.1

Query Match 24.1%; Score 249; DB 2; Length 134;
Best Local Similarity 40.3%; Pred. No. 3.3e-15;
Matches 52; Conservative 28; Mismatches 43; Indels 6; Gaps 4;

QY 55 CRELHYRFLTDGCRSAKPVTELVCSCGCGPARLLPNAIG---RVKWW-RPNGPDFRCI 110
DB 3 CRELSTKYISDGQCTSI SPLKELVCAGECLLLPVLPNWIGGYGTYKWSRSSGEWCV 62
QY 111 PDYRAQRVQLLCPGGAAPRRKRVLRVASCCKRLTRFHNQSELKDFGFPETARP-QKGRK 169
DB 63 NDKTRTQRIQLCCQDGST-RTYKTIIVTACKCKRYTRQHNESHNFESMSPAKPVQHRE 121

QY 170 PRPGAKANQ 178
DB 122 RKRASKSSK 130

RESULT 2

PNO141
lutropin beta chain - sperm whale

N:Alternate names: luteinizing hormone beta chain

C:Species: Physeter catodon (sperm whale)

C>Date: 07-May-1993 #sequence_revision 07-May-1993 #text_change 07-May-1999

C:Accession: PNO141

R:Pankov, Y.A.; Karasev, V.S.

Biokhimiia 49, 1004-1018, 1984

A>Title: Luteinizing hormone of the sperm-whale: amino acid sequence of reduced and carb

A:Reference number: PNO141; MUID:84281133; PMID:6466737

A:Accession: PNO141

A:Molecule type: protein

A:Residues: 1-118 <PAM>

A:Note: article in Russian with English abstract

C:Superfamily: pituitary glycoprotein hormone beta chain

C:Keywords: glycoprotein; hormone

F;9-34,23-57,26-88,38-110,72-100,90-93/Disulfide bonds: #status predicted

KW BEER; gene therapy; antisense therapy; fracture; bone mineralization.
XX
OS Bos taurus.
XX
PN WO200032773-A1.
XX
PD 08-JUN-2000.
XX
PF 24-NOV-1999; 99WO-US27990.
XX
PR 27-NOV-1998; 98US-0110283.
XX
PA (DARW-) DARWIN DISCOVERY LTD.
XX
PI Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepers BW;
PI Van Ness J, Winkler DG;
XX
DR WPI; 2000-412321/35.
XX N-FSDB; AAA29060.
XX
PT Nucleic acids (-) encoding a transforming growth factor beta binding
PT protein, useful for identifying agents for treating osteoporosis,
PT osteoporosis and fractures
XX
PS Claim 7; Page 127; 162pp; English.
XX
CC This shows a bovine transforming growth factor-beta (TGF-beta)
CC binding protein designated bBEER. The cDNA and protein may be used for
CC prevention, treatment and diagnosis of diseases associated with
CC inappropriate BEER expression. For example, they may be used to treat
CC disorders associated with decreased TGF-beta BP expression. The cDNA or
CC vectors may be administered to treat diseases by rectifying mutations or
CC deletions in a patient's genome that affect the activity of BEER by
CC expressing inactive proteins or to supplement the patient's own production
CC of BEER polypeptides. The nucleic acids may be used for recombinant
CC production of BEER, gene therapy, antisense therapy, as probes for
CC diagnostic assays and for functional studies. BEER may be used to raise
CC antibodies and for identification of BEER modulators. BEER antagonists
CC may be used to increase bone mineral content for the treatment of
CC disorders such as osteoporosis, osteoporosis, fractures and other disorders
CC associated with low mineral content.
XX
SQ Sequence 176 AA;

Query Match 81.3%; Score 839; DB 21; Length 176;
Best Local Similarity 86.4%; Pred. No. 1.8e-70;
Matches 152; Conservative 11; Mismatches 9; Indels 4; Gaps 2;

QY 8 NDATVPIGLGEYEPPE-NNQTNRAENGGRPPHPPYDAXDVSEYSCRELHYTRFLTD 66
Db ||||:||||||| ||||:||||||| ||||:||||||| ||||:||||||| ||||:|||||||
1 NDATVPIGLGEYEPPE-NNQTNRAENGGRPPHPPYDAXDVSEYSCRELHYTRFLTD 60
QY 67 GPCRSAPVTCLVSCGCGPALLPNAIGRWKWRPNPNDPFCIPDRYRAQVQLLCFPG 126
Db ||||:||||||| ||||:||||||| ||||:||||||| ||||:||||||| ||||:|||||||
61 GPCRSAPVTCLVSCGCGPALLPNAIGRWKWRPNPNDPFCIPDRYRAQVQLLCFPG 120
QY 127 AAPRSKRVLVASCKCKRLTRFHNSQLKDFGPETARPKGRKRP---GAKANQA 179
Db ||||:||||||| ||||:||||||| ||||:||||||| ||||:||||||| ||||:|||||||
121 AAPRSKRVLVASCKCKRLTRFHNSQLKDFGPETARPKGRKRP---GAKANQA 176

Search completed: March 28, 2003, 14:16:14
Job time : 52.8 secs

XX WPI; 2000-412321/35.
DR N-PSDB; AAA29056.
XX Nucleic acids (I) encoding a transforming growth factor beta binding
PT protein, useful for identifying agents for treating osteopenia,
PT osteoporosis and fractures
XX
XX Claim 3; Page 119-120; 162pp; English.
XX This shows a variant human transforming growth factor-beta (TGF-beta)
CC binding protein designated BEER V101, which comprises a substitution of
CC isoleucine for the wild-type valine at residue 10. The cDNA and protein
CC may be used for prevention, treatment and diagnosis of diseases
CC associated with inappropriate BEER expression. For example, they may be
CC used to treat disorders associated with decreased TGF-beta BP expression.
CC The cDNA or vectors may be administered to treat diseases by rectifying
CC mutations or deletions in a patient's genome that affect the activity of
CC BEER by expressing inactive proteins or to supplement the patients own
CC production of BEER polypeptides. The nucleic acids may be used for
CC recombinant production of BEER, gene therapy, antisense therapy, as
CC probes for diagnostic assays and for functional studies. BEER may be used
CC to raise antibodies and for identification of BEER modulators. BEER
CC antagonists may be used to increase bone mineral content for the
CC treatment of disorders such as osteopenia, osteoporosis, fractures and
CC other disorders associated with low mineral content.
XX
SQ Sequence 213 AA;
Query Match 89.5%; Score 923.5; DB 21; Length 213;
Best Local Similarity 87.9%; Pred. No. 2.9e-78;
Matches 167; Conservative 10; Mismatches 8; Indels 5; Gaps 2;
OY 1 QGWAQFRNDATVPIGLGEYEPPEPP--ENNQTMNRAENGRRPHHPYDAKDVSEYSCREL 58
Db 24 QGWAQFRNDATVPIGLGEYEPPEPP--ENNQTMNRAENGRRPHHPYDAKDVSEYSCREL 83
OY 59 HVTRELTGDCRSAPKPVTELVCSCGCGPARLLPNAIGRKWWRPNPNDPFCIPDRYRAQR 118
Db 84 HFTRYVTGDCRSAPKPVTELVCSCGCGPARLLPNAIGRKWWRPNPNDPFCIPDRYRAQR 143
OY 119 VQLCPGGAAPRRKRVRLVASCCKRLTRFHNQSELKDFGTEARPOKGRKPRP---GAK 175
Db 144 VQLCPGGAAPRRKRVRLVASCCKRLTRFHNQSELKDFGTEARPOKGRKPRPARSAK 203
OY 176 ANQAELENAY 185
Db 204 ANQAELENAY 213
RESULT 14
AA96436
ID AA96436 standard; Protein; 213 AA.
XX
XX AC AA96436;
XX 12-SEP-2000 (first entry)
XX Human TGF-beta binding protein (BEER) variant P38R.
XX osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
XX BEER; variant; P38R; gene therapy; antisense therapy; fracture;
XX chromosome 17q12-21; bone mineralization.
XX Homo sapiens.
XX Key Location/Qualifiers
FH Misc-difference 38
FT /label= P38R
FT /note= "wild type proline has been substituted with
FT arginine"
XX
XX W0200032773-A1.
PN

XX 08-JUN-2000.
XX 24-NOV-1999; 99WO-US27990.
XX 27-NOV-1998; 98US-0110283.
XX (DARW-) DARWIN DISCOVERY LTD.
XX Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepers BW;
PI Van Ness J, Winkler DG;
XX WPI; 2000-412321/35.
DR N-PSDB; AAA29062.
XX Nucleic acids (I) encoding a transforming growth factor beta binding
PT protein, useful for identifying agents for treating osteopenia,
PT osteoporosis and fractures
XX
XX Disclosure; Page 121; 162pp; English.
XX This shows a variant human transforming growth factor-beta
CC (TGF-beta) binding protein designated BEER P38R. The encoded protein
CC comprises a substitution of arginine for the wild-type proline at
CC residue 38. The cDNA and protein may be used for prevention, treatment
CC and diagnosis of diseases associated with inappropriate BEER expression.
CC For example, they may be used to treat disorders associated with
CC decreased TGF-beta BP expression. The cDNA or vectors may be administered
CC to treat diseases by rectifying mutations or deletions in a patient's
CC genome that affect the activity of BEER by expressing inactive proteins
CC or to supplement the patients own production of BEER polypeptides. The
CC functional studies. BEER may be used to raise antibodies and for
CC therapy, antisense therapy, as probes for diagnostic assays and for
CC identification of BEER modulators. BEER antagonists may be used to
CC increase bone mineral content for the treatment of disorders such as
CC osteopenia, osteoporosis, fractures and other disorders associated with
CC low mineral content.
XX
SQ Sequence 213 AA;
Query Match 89.5%; Score 923.5; DB 21; Length 213;
Best Local Similarity 87.9%; Pred. No. 2.9e-78;
Matches 167; Conservative 10; Mismatches 8; Indels 5; Gaps 2;
OY 1 QGWAQFRNDATVPIGLGEYEPPEPP--ENNQTMNRAENGRRPHHPYDAKDVSEYSCREL 58
Db 24 QGWAQFRNDATVPIGLGEYEPPEPP--ENNQTMNRAENGRRPHHPYDAKDVSEYSCREL 83
OY 59 HVTRELTGDCRSAPKPVTELVCSCGCGPARLLPNAIGRKWWRPNPNDPFCIPDRYRAQR 118
Db 84 HFTRYVTGDCRSAPKPVTELVCSCGCGPARLLPNAIGRKWWRPNPNDPFCIPDRYRAQR 143
OY 119 VQLCPGGAAPRRKRVRLVASCCKRLTRFHNQSELKDFGTEARPOKGRKPRP---GAK 175
Db 144 VQLCPGGAAPRRKRVRLVASCCKRLTRFHNQSELKDFGTEARPOKGRKPRPARSAK 203
OY 176 ANQAELENAY 185
Db 204 ANQAELENAY 213
RESULT 15
AA96434
ID AA96434 standard; Protein; 176 AA.
XX
XX AC AA96434;
XX 12-SEP-2000 (first entry)
XX Bovine TGF-beta binding protein (BEER).
XX osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
XX

Query Match 90.4%; Score 932.5; DB 23; Length 213;
 Best Local Similarity 88.4%; Pred. No. 4.2e-79;
 Matches 168; Conservative 10; Mismatches 7; Indels 5; Gaps 2;

QY 1 QGWAQFRNDATVPGIGYPPPPP--ENNQTMNRAENGRRPHHPYDAKDVSEYSCREL 58
 DB 24 QGWAQFRNDATVPGIGYPPPPP--ENNQTMNRAENGRRPHHPYDAKDVSEYSCREL 83
 QY 59 HYTRFLTDGPCRSAPVTELVCSCGQGPAPLLPNAIGRVKWRPNPDPFCIPDRYRAOR 118
 DB 84 HTRYVTDGPCRSAPVTELVCSCGQGPAPLLPNAIGRVKWRPNPDPFCIPDRYRAOR 143
 QY 119 VOLLCPGGAAPRSKRVRLVASCKKRLTRFHNQSELKDFGPTARPOKGRKPRP---GAK 175
 DB 144 VOLLCPGGAAPRSKRVRLVASCKKRLTRFHNQSELKDFGPTARPOKGRKPRP---GAK 203
 QY 176 ANQAELENAY 185
 DB 204 ANQAELENAY 213

RESULT 12
 AAEL17089
 ID AAEL17089 standard; Protein; 213 AA.
 AC AAEL17089;
 DT 18-APR-2002 (first entry)
 DE Human osteolevin protein.
 KW Human; osteolevin; osteopathic; cytostatic; bone formation; osteoporosis;
 KW Van Buchem-sclerosteosis disease; sclerosteosis; transgenic animal;
 KW Paget's disease; chromosome 17.
 OS Homo sapiens.
 FH Key
 FT Peptide 1..19
 FT /label= Signal_peptide
 FT Protein 20..213
 FT /label= Mature_osteolevin_protein
 FT Misc-difference 10
 FT /note= "During polymorphism wild type Val is substituted with Ile"
 XX WO200198491-A2.
 XX 27-DEC-2001.
 XX 15-JUN-2001; 2001WO-EP06795.
 XX 19-JUN-2000; 2000EP-0112867.
 XX (HOFF) HOFFMANN LA ROCHE & CO AG F.
 XX (UYIN-) UNIV INSTELLING ANTWERPEN UIA.
 XX Balenans W, Ebeling M, Foerzler D, Patel N, Van Hul W;
 XX Vickery BH;
 XX WPI; 2002-139789/18.
 XX N-PSDB; AAD27576, AAD27577.
 XX Novel genetic polymorphisms in the Van Buchem-sclerosteosis disease
 XX region that are associated with abnormal bone formation useful for
 XX diagnosis and assessment of osteoporosis or sclerosteosis in humans
 XX Claim 11; Fig 4; 70pp; English.
 XX The invention relates to a nucleic acid encoding osteolevin region
 XX polymorphisms. The invention also relates to genetic polymorphisms in
 XX the Van Buchem-sclerosteosis disease region that are associated with
 XX disorders resulting in either net excess bone formation or insufficient

CC bone formation in humans. Osteolevin DNA is useful for screening for
 CC osteolevin polymorphisms associated with abnormal bone formation in a
 CC subject and for the presence of a heritably linked form of abnormal bone
 CC formation in a subject, by determining the presence of a polymorphism in
 CC the osteolevin nucleic acid sequence obtained from the subject.
 CC Osteolevin protein is useful for treating diseases associated with
 CC abnormal bone formation, such as sclerosteosis, Van Buchem's disease and
 CC Paget's disease. Nucleic acids which encode genes in the osteolevin
 CC region or their modified forms can also be used to generate either
 CC transgenic animals or knockout animals which are useful in the screening
 CC and development of therapeutically useful reagents. Osteolevin proteins
 CC are useful in pharmacological characterisation of novel modulators of the
 CC activity of protein and protein complexes. Human osteolevin gene is
 CC located on chromosome 17. The present sequence is human osteolevin
 CC protein.
 XX
 SQ Sequence 213 AA;

Query Match 90.4%; Score 932.5; DB 23; Length 213;
 Best Local Similarity 88.4%; Pred. No. 4.2e-79;
 Matches 168; Conservative 10; Mismatches 7; Indels 5; Gaps 2;

QY 1 QGWAQFRNDATVPGIGYPPPPP--ENNQTMNRAENGRRPHHPYDAKDVSEYSCREL 58
 DB 24 QGWAQFRNDATVPGIGYPPPPP--ENNQTMNRAENGRRPHHPYDAKDVSEYSCREL 83
 QY 59 HYTRFLTDGPCRSAPVTELVCSCGQGPAPLLPNAIGRVKWRPNPDPFCIPDRYRAOR 118
 DB 84 HTRYVTDGPCRSAPVTELVCSCGQGPAPLLPNAIGRVKWRPNPDPFCIPDRYRAOR 143
 QY 119 VOLLCPGGAAPRSKRVRLVASCKKRLTRFHNQSELKDFGPTARPOKGRKPRP---GAK 175
 DB 144 VOLLCPGGAAPRSKRVRLVASCKKRLTRFHNQSELKDFGPTARPOKGRKPRP---GAK 203
 QY 176 ANQAELENAY 185
 DB 204 ANQAELENAY 213

RESULT 13
 AAY96430
 ID AAY96430 standard; Protein; 213 AA.
 AC AAY96430;
 DT 12-SEP-2000 (first entry)
 DE Human TGF-beta binding protein (BEER) variant V101.
 DE osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
 DE BEER; variant; V101; gene therapy; anticense therapy; fracture;
 DE chromosome 17q12-21; bone mineralization.
 OS Homo sapiens.
 FH Key
 FT Location/Qualifiers
 FT Misc-difference 10
 FT /label= V101
 FT /note= "wild type valine has been substituted with
 FT isoleucine"

WO200032773-A1.
 08-JUN-2000.
 24-NOV-1999; 99WO-US27990.
 27-NOV-1998; 98US-0110283.
 (DARW-) DARWIN DISCOVERY LTD.
 Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepers BW;
 Van Ness J, Winkler IG;

DB		204 ANQAELENAY 213
	RESULT 11	
	ABB07209	
ID	ABB07209 standard; Protein; 213 AA.	
XX	XX	
AC	ABB07209;	
XX	XX	
DT	26-MAR-2002 (first entry)	
XX	XX	
DE	Human cloaked-2 polypeptide sequence.	
XX	XX	
KW	Cloaked-2; cysteine knot motif; nephrotropic; cardiant; immunomodulator;	
KW	hepatotropic; antiinflammatory; antithyroid; cytostatic; neuroprotective;	
KW	antianemic; hypotensive; antiarrhythmic; antiateriosclerotic; muscular;	
KW	antidiabetic; anorectic; gene therapy; cell therapy; antisense therapy;	
XX	human.	
XX	XX	
OS	Homo sapiens.	
XX	XX	
FH	Key	Location/Qualifiers
FT	Peptide	1..23
FT	/note= "signal peptide"	
FT	Protein	24..213
FT	/note= "mature protein (ABB07207)"	
XX	XX	
FN	WO200192308-A2.	
XX	XX	
PD	06-DEC-2001.	
XX	XX	
PF	29-MAY-2001; 2001WO-US17478.	
XX	XX	
PR	01-JUN-2000; 2000US-208550P.	
PR	04-AUG-2000; 2000US-223542P.	
XX	XX	
PA	(AMGE-) AMGEN INC.	
PI	Paszty CJ, Gao Y;	
DR	WPI; 2002-114325/15.	
DR	N-PSDB; ABA94293.	
PT	New human and mouse cystine-knot polypeptide designated as Cloaked-2,	
PT	for treating or preventing kidney, heart (e.g. myocardial infarction)	
PT	or liver (e.g. hepatitis) diseases -	
XX	XX	
FS	Example 1; Fig 1; 170pp; English.	
XX	XX	
CC	The invention relates to polypeptides comprising a cysteine knot motif	
CC	and designated as Cloaked-2, derived from human and mouse. The cloaked-2	
CC	polypeptides can be expressed by standard recombinant methodology. The	
CC	cloaked-2 polynucleotides are useful in gene therapy and antisense	
CC	therapy. The cloaked-2 polypeptides and polynucleotides are useful for	
CC	treating, preventing, ameliorating or detecting diseases and disorders of	
CC	the kidney (e.g. anemia, hypertension or low blood pressure), heart (e.g.	
CC	cardiac hypertrophy, congestive heart failure, myocardial infarction,	
CC	rhythmias, atherosclerosis, hypertension or low blood pressure), placenta (e.g.	
CC	skeletal muscle (e.g. muscular dystrophy or cachexia), liver (e.g. hepatitis or	
CC	congenital abnormalities (e.g. diabetes or pancreatitis), thyroid (e.g.	
CC	cirrhosis), pancreas (e.g. diabetes or adrenal cortex (e.g. Cushing's disease	
CC	grave's disease or myxedema) or adrenal cortex (e.g. Cushing's disease	
CC	or Addison's disease), homeostasis or metabolic diseases (e.g. obesity,	
CC	cancer or myopathies), infections, or autoimmune diseases. Selective	
CC	binding agents may be used to modulate the biological activities of	
CC	cloaked-2 polypeptides or to detect cloaked-2 polypeptide levels in a	
CC	sample. Transgenic non-human animals are useful for drug candidate	
CC	screening. The present sequence represents the human cloaked-2	
XX	XX	
SQ	Sequence 213 AA;	

CC Grave's disease or myxedema) or adrenal cortex (e.g. Cushing's disease
 CC or Addison's disease), homeostasis or metabolic diseases (e.g. obesity,
 CC cancer or myopathies), infections, or autoimmune diseases. Selective
 CC binding agents may be used to modulate the biological activities of
 CC Cloaked-2 polypeptides or to detect Cloaked-2 polypeptide levels in a
 CC sample. Transgenic non-human animals are useful for drug candidate
 CC screening. The present sequence represents the human cloaked-2
 CC mature polypeptide sequence.

XX Sequence 190 AA;
 SQ Query Match 90.4%; Score 932.5; DB 23; Length 190;
 Best Local Similarity 88.4%; Pred. No. 3.7e-79;
 Matches 168; Conservative 10; Mismatches 7; Indels 5; Gaps 2;
 QY 1 QGWAQFNDATEVIGLGEYEPPEPP--ENNQTMNRAENGRRPHHPYDAKDVSEYSCREL 58
 DB 1 QGWAQFNDATEVIGLGEYEPPEPP--ENNQTMNRAENGRRPHHPYDAKDVSEYSCREL 60
 QY 59 HYTRFLTDGPCRSAPVTELVCSGQCGPARLLPNAIGRVKWRPNPGDFRCIPDRYRAQR 118
 DB 61 HFTRYVTDGPCRSAPVTELVCSGQCGPARLLPNAIGRVKWRPNPGDFRCIPDRYRAQR 120
 QY 119 VQLLCPGGAAPRSRKVRLVASCKCKRLTRFHNQSELKDFGTEAARPKGKRP--GAK 175
 DB 121 VQLLCPGGAAPRSRKVRLVASCKCKRLTRFHNQSELKDFGTEAARPKGKRP--GAK 180
 QY 176 ANQAELENAY 185
 DB 181 ANQAELENAY 190

RESULT 7
 AAB26106
 ID AAB26106 standard; Protein; 213 AA.
 XX AC AAB26106;
 XX DT 15-JAN-2001 (first entry)
 XX DE Human DAN/Cerberus-related protein 6 (hDCR6) (exons 1 and 4).
 XX KW Human; DNA/Cerberus-related protein 6; hDCR6; morphogenic protein;
 XX KW antagonist; BMP; cell growth; cell differentiation; bone formation;
 XX KW gene therapy.
 XX OS Homo sapiens.
 XX PN WO200055193-A2.
 XX PD 21-SEP-2000.
 XX PF 02-MAR-2000; 2000WO-US05537.
 XX PR 12-MAR-1999; 99US-0124118.
 XX PA (REG-) REGENERON PHARM INC.
 XX PI Economides AN;
 XX DR WPI; 2000-638179/61.
 XX DR N-PSDB; AAA94051.
 XX PT Novel isolated, human DNA/Cerberus related protein 6 which include
 XX PT natural homologue, and polypeptides comprising DCR6 domain and nucleic
 XX PT acids encoding the proteins which are useful as probes and primers -
 XX PS Claim 8; Fig 3; 40pp; English.
 XX CC The present sequence comprises the amino acid sequence encoded by exons 1
 CC and 4 of the human DAN/Cerberus-related protein 6 (hDCR6) coding
 CC sequence. The coding sequence was isolated from a human kidney cDNA
 CC library containing exons 1 and 4 of the sequence. hDCR6 is closely

CC related to the DAN and DCR5 proteins, both of which act as antagonists of
 CC morphogenic proteins such as BMP. It is possible that the hDCR6 gene and
 CC protein can be used as immunogens, modulators of cell function, growth
 CC and differentiation, to reduce undesirable bone formation, to identify
 CC DCR6 binding agents, in diagnosis, and in gene therapy.

XX Sequence 213 AA;
 SQ Query Match 90.4%; Score 932.5; DB 21; Length 213;
 Best Local Similarity 88.4%; Pred. No. 4.2e-79;
 Matches 168; Conservative 10; Mismatches 7; Indels 5; Gaps 2;
 QY 1 QGWAQFNDATEVIGLGEYEPPEPP--ENNQTMNRAENGRRPHHPYDAKDVSEYSCREL 58
 DB 24 QGWAQFNDATEVIGLGEYEPPEPP--ENNQTMNRAENGRRPHHPYDAKDVSEYSCREL 83
 QY 59 HYTRFLTDGPCRSAPVTELVCSGQCGPARLLPNAIGRVKWRPNPGDFRCIPDRYRAQR 118
 DB 84 HFTRYVTDGPCRSAPVTELVCSGQCGPARLLPNAIGRVKWRPNPGDFRCIPDRYRAQR 143
 QY 119 VQLLCPGGAAPRSRKVRLVASCKCKRLTRFHNQSELKDFGTEAARPKGKRP--GAK 175
 DB 144 VQLLCPGGAAPRSRKVRLVASCKCKRLTRFHNQSELKDFGTEAARPKGKRP--GAK 203
 QY 176 ANQAELENAY 185
 DB 204 ANQAELENAY 213

RESULT 8
 AAY96429
 ID AAY96429 standard; Protein; 213 AA.
 XX AC AAY96429;
 XX DT 12-SEP-2000 (first entry)
 XX DE Human TGF-beta binding protein (BEER).
 XX KW osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
 XX KW BEER; gene therapy; antisense therapy; fracture; bone mineralization.
 XX OS Homo sapiens.
 XX PN WO200032773-A1.
 XX PD 08-JUN-2000.
 XX PF 24-NOV-1999; 99WO-US27990.
 XX PR 27-NOV-1998; 98US-0110283.
 XX PA (DARW-) DARWIN DISCOVERY LTD.
 XX PI Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepers BW;
 XX PI Van Ness J, Winkler DG;
 XX XX WPI; 2000-412321/35.
 XX DR N-PSDB; AAA29055.
 XX PT Nucleic acids (1) encoding a transforming growth factor beta binding
 XX PT protein, useful for identifying agents for treating osteopenia,
 XX PT osteoporosis and fractures
 XX PS Claim 2; Page 116; 162pp; English.
 XX CC This shows the human transforming growth factor-beta (TGF-beta)
 CC binding protein designated hBEER. The cDNA and protein may be used for
 CC prevention, treatment and diagnosis of diseases associated with
 CC inappropriate BEER expression. For example, they may be used to treat
 CC disorders associated with decreased TGF-beta BP expression. The cDNA or
 CC vectors may be administered to treat diseases by rectifying mutations or
 CC deletions in a patient's genome that affect the activity of BEER by

Db 84 HYTRFVTDGFCRSAPKPVTELVCSCGCGPARLLPNAIGRKWWRNGPDPRCIPDRYRAQR 143
 QY 119 VOLLCPGGAAPRSRKVRLVASCCKRLTRFNQSELKDFGPETARPOKGRKPRP---GAK 175
 Db 144 VOLLCPGGAAPRSRKVRLVASCCKRLTRFNQSELKDFGPETARPOKGRKPRPARGAK 203
 QY 176 ANQAELENAY 185
 Db 204 ANQAELENAY 213

RESULT 5
 AAY96431
 ID AAY96431 standard; Protein; 213 AA.
 XX
 AC AAY96431;
 DT 12-SEP-2000 (first entry)
 DE Vervet TGF-beta binding protein (BEER).
 KW osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
 KW BEER; gene therapy; antisense therapy; fracture; bone mineralization.
 XX Cercopithecus pygerythrus.
 OS
 PN W0200032773-A1.
 PD 08-JUN-2000.
 XX
 PF 24-NOV-1999; 99WO-US27990.
 XX
 PR 27-NOV-1998; 98US-0110283.
 XX
 PA (DARW-) DARWIN DISCOVERY LTD.
 PI Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepers BW;
 PI Van Ness J, Winkler DG;
 XX
 DR WPI; 2000-412321/35.
 DR N-PSDB; AAA29057.
 XX

Nucleic acids (1) encoding a transforming growth factor beta binding protein, useful for identifying agents for treating osteopenia, osteoporosis and fractures
 Claim 4; Page 122-123; 162pp; English.

This shows a vervet transforming growth factor-beta (TGF-beta) binding protein designated vBEER. The cDNA and protein may be used for prevention, treatment and diagnosis of diseases associated with inappropriate BEER expression. For example, they may be used to treat disorders associated with decreased TGF-beta BP expression. The cDNA or vectors may be administered to treat diseases by rectifying mutations or deletions in a patient's genome that affect the activity of BEER by expressing inactive proteins or to supplement the patients own production of BEER polypeptides. The nucleic acids may be used for recombinant production of BEER, gene therapy, antisense therapy, as probes for diagnostic assays and for functional studies. BEER may be used to raise antibodies and for identification of BEER modulators. BEER antagonists may be used to increase bone mineral content for the treatment of disorders such as osteopenia, osteoporosis, fractures and other disorders associated with low mineral content.

XX Sequence 213 AA;
 SQ

Query Match
 Best Local Similarity 92.2%; Score 951.5; DB 21; Length 213;
 Matches 171; Conservative 10; Mismatches 4; Indels 5; Gaps 2;
 QY 1 QGWAQFRNDATVIGLGYRPPPP--ENNQTWNAENGRRPHHPYDAKDVSEYSREL 58
 Db 24 QGWAQFRNDATVIGLGYRPPPPLENNKTWNAENGRRPHHPYDAKDVSEYSREL 83

QY 59 HYTRFVTDGFCRSAPKPVTELVCSCGCGPARLLPNAIGRKWWRNGPDPRCIPDRYRAQR 118
 Db 84 HFRYVTDGFCRSAPKPVTELVCSCGCGPARLLPNAIGRKWWRNGPDPRCIPDRYRAQR 143
 QY 119 VOLLCPGGAAPRSRKVRLVASCCKRLTRFNQSELKDFGPETARPOKGRKPRP---GAK 175
 Db 144 VOLLCPGGAAPRSRKVRLVASCCKRLTRFNQSELKDFGPETARPOKGRKPRPARGAK 203
 QY 176 ANQAELENAY 185
 Db 204 ANQAELENAY 213

RESULT 6
 ABB07207
 ID ABB07207 standard; Protein; 190 AA.
 XX
 AC ABB07207;
 DT 26-MAR-2002 (first entry)
 DE Human cloaked-2 polypeptide mature protein sequence.
 XX
 KW Cloaked-2; cysteine knot motif; nephrotropic; cardiant; immunomodulator;
 KW hepatotropic; antiinflammatory; antithyroid; cytostatic; neuroprotective;
 KW antianemic; hypotensive; antiarrhythmic; antiarteriosclerotic; muscular;
 KW antidiabetic; anorectic; gene therapy; cell therapy; antisense therapy;
 XX human.
 OS Homo sapiens.
 FH
 FT Key Location/Qualifiers
 FT Misc-difference 1.16
 FT /note= "this region has been repeated twice in the
 FT sequence provided in the sequence listing but has
 FT been indicated correctly in the sequence in the
 FT Figure"
 XX
 PN W0200192308-A2.
 XX
 PD 06-DEC-2001.
 XX
 PF 29-MAY-2001; 2001WO-US17478.
 XX
 PR 01-JUN-2000; 2000US-208550P.
 PR 04-AUG-2000; 2000US-223542P.
 XX
 PA (AMGE-) AMGEN INC.
 XX
 PI Paszty CJ, Gao Y;
 XX
 DR WPI; 2002-114325/15.
 DR N-PSDB; ABA94293.
 XX

New human and mouse cysteine-knot polypeptide designated as Cloaked-2, for treating or preventing kidney, heart (e.g. myocardial infarction) or liver (e.g. hepatitis) diseases
 Claim 13; Fig 1; 170pp; English.

The invention relates to polypeptides comprising a cysteine knot motif and designated as Cloaked-2, derived from human and mouse. The cloaked-2 polypeptides can be expressed by standard recombinant methodology. The cloaked-2 polynucleotides are useful in gene therapy and antisense therapy. The cloaked-2 polypeptides and polynucleotides are useful for treating, preventing, ameliorating or detecting diseases and disorders of the kidney (e.g. anemia, hypertension or low blood pressure), heart (e.g. cardiac hypertrophy, congestive heart failure, myocardial infarction, arrhythmias, atherosclerosis, hypertension or low blood pressure), skeletal muscle (e.g. muscular dystrophy or cachexia), placenta (e.g. congenital abnormalities or miscarriage), liver (e.g. hepatitis or cirrhosis), pancreas (e.g. diabetes or pancreatitis), thyroid (e.g.

Mon Mar 31 09:44:35 2003

antidiabetic; anorectic; gene therapy; cell therapy; antisense therapy; mouse.

Mus musculus.

Key Location/Qualifiers

Peptide 1..23 /note= "signal peptide"

Protein 24 /note= "mature protein (AB07208)"

Misc-difference 198..200 /note= "this region is missing in the sequence provided in the sequence listing but has been indicated correctly in the sequence in the Figure"

WO200192308-A2.

06-DEC-2001.

29-MAY-2001; 2001WO-US17478.

01-JUN-2000; 2000US-208550P.

04-AUG-2000; 2000US-223542P.

(AMGE-) AMGEN INC.

Paszty CJ, Gao Y;

WPI; 2002-114325/15.

N-PSDB; ABA94294.

New human and mouse cysteine-knot polypeptide designated as Cloaked-2, for treating or preventing kidney, heart (e.g. myocardial infarction) or liver (e.g. hepatitis) diseases

Example 2; Fig 2; 170pp; English.

The invention relates to polypeptides comprising a cysteine knot motif and designated as Cloaked-2, derived from human and mouse. The cloaked-2 polypeptides can be expressed by standard recombinant methodology. The cloaked-2 polynucleotides are useful in gene therapy and antisense therapy. The cloaked-2 polypeptides and polynucleotides are useful for treating, preventing or detecting diseases and disorders of the kidney (e.g. anemia, hypertension or low blood pressure), heart (e.g. cardiac hypertrophy, congestive heart failure, myocardial infarction, arrhythmias, atherosclerosis, hypertension or low blood pressure), skeletal muscle (e.g. muscular dystrophy or cachexia), placenta (e.g. cirrhosis), pancreas (e.g. diabetes or pancreatitis), liver (e.g. hepatitis or Grave's disease or myxedema) or adrenal cortex (e.g. Cushing's disease or Addison's disease), homeostasis or metabolic diseases (e.g. obesity, cancer or myopathies), infections, or autoimmune diseases. Selective binding agents may be used to modulate the biological activities of Cloaked-2 polypeptides or to detect Cloaked-2 polypeptide levels in a sample. Transgenic non-human animals are useful for drug candidate screening. The present sequence represents the mouse cloaked-2 polypeptide.

Sequence 211 AA;

Query Match 98.9%; Score 1020.5; DB 23; Length 211;

Best Local Similarity 98.4%; Pred. No. 2.5e-87; Indels 3; Gaps 1;

Matches 185; Conservative 0; Mismatches 0;

1 QGQAFRNDATEVIFGLGEYPPPPNNQTNRAENGRRPHHPHYDAKDVSEYSCRELHY 60

24 QGQAFRNDATEVIFGLGEYPPPPNNQTNRAENGRRPHHPHYDAKDVSEYSCRELHY 83

61 TRFLTDGCRSAKPVTELVCSQCQGPALLPNAIGRVKWRPNPDPFCIPDRYRAQRVQ 120

84 TRFLTDGCRSAKPVTELVCSQCQGPALLPNAIGRVKWRPNPDPFCIPDRYRAQRVQ 143

121 LLCFGGAAPRSRKVLVASCCKRLTRFHNSQLKDFGPETARFQKGRKPRP---GAKAN 177

Db 144 LLCFGGAAPRSRKVLVASCCKRLTRFHNSQLKDFGPETARFQKGRKPRGARGAKAN 203

Qy 178 QAELENAY 185

Db 204 QAELENAY 211

RESULT 4
ID AAY96433 standard; Protein; 213 AA.
XX AAY96433;
XX 12-SEP-2000 (first entry)
XX Rat TGF-beta binding protein (BEER).
XX osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
XX BEER; gene therapy; antisense therapy; fracture; bone mineralization.
XX Rattus norvegicus.
XX WO200032773-A1.
XX 08-JUN-2000.
XX 24-NOV-1999; 99WO-US27990.
XX 27-NOV-1998; 98US-0110283.
XX (DARW-) DARWIN DISCOVERY LTD.
XX Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepers BW;
XX Van Ness J, Winkler DG;
XX WPI; 2000-412321/35.
XX N-PSDB; AAY29059.

Nucleic acids (I) encoding a transforming growth factor beta binding protein, useful for identifying agents for treating osteoporosis, osteoporosis and fractures

Claim 6; Page 125-126; 162pp; English.

This shows a rat transforming growth factor-beta (TGF-beta) binding protein designated rBEER. The CDNA and protein may be used for prevention, treatment and diagnosis of diseases associated with inappropriate BEER expression. For example, they may be used to treat disorders associated with decreased TGF-beta BP expression. The CDNA or vectors may be administered to treat diseases by rectifying mutations or deletions in a patient's genome that affect the activity of BEER by expressing inactive proteins or to supplement the patient's own production of BEER polypeptides. The nucleic acids may be used for recombinant production of BEER, gene therapy, antisense therapy, as probes for diagnostic assays and for functional studies. BEER may be used to raise antibodies and for identification of BEER modulators. BEER antagonists may be used to increase bone mineral content for the treatment of disorders such as osteoporosis, osteoporosis, fractures and other disorders associated with low mineral content.

Sequence 213 AA;

Query Match 95.2%; Score 982.5; DB 21; Length 213;

Best Local Similarity 94.2%; Pred. No. 9e-84; 3; Indels 5; Gaps 2;

Matches 179; Conservative 3; Mismatches 3;

1 QGQAFRNDATEVIFGLGEYPPPP--ENNQTNRAENGRRPHHPHYDAKDVSEYSCREL 58

24 QGQAFRNDATEVIFGLGEYPPPPQLENNQTNRAENGRRPHHPHYDTKDVSEYSCREL 83

59 HYTRFLTDGCRSAKPVTELVCSQCQGPALLPNAIGRVKWRPNPDPFCIPDRYRAQR 118

(AMGE-) AMGEN INC.

Paszy CJ, Gao Y;

WPI: 2002-114325/15.

N-PSDB; ABA94294.

New human and mouse cysteine-knot polypeptide designated as Cloaked-2, PT
PT for treating or preventing kidney, heart (e.g. myocardial infarction)
or liver (e.g. hepatitis) diseases

Claim 13; Fig 2; 170pp; English.

The invention relates to polypeptides comprising a cysteine knot motif
and designated as Cloaked-2, derived from human and mouse. The cloaked-2
polypeptides can be expressed by standard recombinant methodology. The
cloaked-2 polynucleotides are useful in gene therapy and antisense
therapy. The cloaked-2 polypeptides and polynucleotides are useful for
treating, preventing, ameliorating or detecting diseases and disorders of
the kidney (e.g. anemia, hypertension or low blood pressure), heart (e.g.
cardiac hypertrophy, congestive heart failure, myocardial infarction,
arrhythmias, atherosclerosis, hypertension or low blood pressure),
skeletal muscle (e.g. muscular dystrophy or cachexia), placenta (e.g.
congenital abnormalities or miscarriage), liver (e.g. hepatitis or
cirrhosis), pancreas (e.g. diabetes or pancreatitis), thyroid (e.g.
Grave's disease or myxedema) or adrenal cortex (e.g. Cushing's disease
or Addison's disease), homeostasis or metabolic diseases (e.g. obesity,
cancer or myopathies), infections, or autoimmune diseases. Selective
binding agents may be used to modulate the biological activities of
Cloaked-2 polypeptides or to detect Cloaked-2 polypeptide levels in a
sample. Transgenic non-human animals are useful for drug candidate
screening. The present sequence represents the mouse cloaked-2 mature
polypeptide.

Sequence 188 AA;

Query Match 98.9%; Score 1020.5; DB 23; Length 188;

Best Local Similarity 98.4%; Pred. No. 2.2e-87;

Matches 185; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

QY 1 QGWAQFRNDATEVIGLGEYPEPPENNQTMRNRAENGRRPHHPYDAKDVSEYSCRELHY 60

DB 1 QGWAQFRNDATEVIGLGEYPEPPENNQTMRNRAENGRRPHHPYDAKDVSEYSCRELHY 60

QY 61 TRFLTDGPCRSKAPVTELVCSGCGCPARLLPNAIGRVKWRPNPGRFCIPDRYRAQRVQ 120

DB 61 TRFLTDGPCRSKAPVTELVCSGCGCPARLLPNAIGRVKWRPNPGRFCIPDRYRAQRVQ 120

QY 121 LLCFGGAAPRSRKVRLVASCCKKRLTRFHNSQSELKDFGPETARPOKGRKPRP---GAKAN 177

DB 121 LLCFGGAAPRSRKVRLVASCCKKRLTRFHNSQSELKDFGPETARPOKGRKPRP---GAKAN 177

QY 178 QAELENAY 185

DB 181 QAELENAY 188

RESULT 2

AA96432

ID AAY96432 standard; Protein; 211 AA.

AC AAY96432;

XX 12-SEP-2000 (first entry)

DT Murine TGF-beta binding protein (BEER).

DE osteopathic; transforming growth factor-beta; TGF-beta; binding protein;

XX BEER; gene therapy; antisense therapy; fracture; bone mineralization.

KW Mus musculus.

XX WO200032773-A1.

PN

XX

PD

XX

PF

XX

PR

XX

PA

XX

PI

XX

PS

XX

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

08-JUN-2000.

24-NOV-1999; 99WO-US27990.

27-NOV-1998; 98US-0110283.

(DAW-) DARWIN DISCOVERY LTD.

Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepers BW;

Van Ness J, Winkler DG;

WPI: 2000-412321/35.

N-PSDB; AAA29058.

Nucleic acids (1) encoding a transforming growth factor beta binding

protein, useful for identifying agents for treating osteopenia,

osteoporosis and fractures

Claim 5; Page 124; 162pp; English.

This shows a murine transforming growth factor-beta (TGF-beta)

binding protein designated mBEER. The cDNA and protein may be used for

prevention, treatment and diagnosis of diseases associated with

inappropriate BEER expression. For example, they may be used to treat

disorders associated with decreased TGF-beta BP expression. The cDNA or

vectors may be administered to treat diseases by rectifying mutations or

deletions in a patient's genome that affect the activity of BEER by

expressing inactive proteins or to supplement the patients own production

of BEER polypeptides. The nucleic acids may be used for recombinant

production of BEER, gene therapy, antisense therapy, as probes for

diagnostic assays and for functional studies. BEER may be used to raise

antibodies and for identification of BEER modulators. BEER antagonists

may be used to increase bone mineral content for the treatment of

disorders such as osteopenia, osteoporosis, fractures and other disorders

associated with low mineral content.

Sequence 211 AA;

Query Match 98.9%; Score 1020.5; DB 21; Length 211;

Best Local Similarity 98.4%; Pred. No. 2.5e-87;

Matches 185; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

QY 1 QGWAQFRNDATEVIGLGEYPEPPENNQTMRNRAENGRRPHHPYDAKDVSEYSCRELHY 60

DB 24 QGWAQFRNDATEVIGLGEYPEPPENNQTMRNRAENGRRPHHPYDAKDVSEYSCRELHY 83

QY 61 TRFLTDGPCRSKAPVTELVCSGCGCPARLLPNAIGRVKWRPNPGRFCIPDRYRAQRVQ 120

DB 84 TRFLTDGPCRSKAPVTELVCSGCGCPARLLPNAIGRVKWRPNPGRFCIPDRYRAQRVQ 143

QY 121 LLCFGGAAPRSRKVRLVASCCKKRLTRFHNSQSELKDFGPETARPOKGRKPRP---GAKAN 177

DB 144 LLCFGGAAPRSRKVRLVASCCKKRLTRFHNSQSELKDFGPETARPOKGRKPRP---GAKAN 203

QY 178 QAELENAY 185

DB 204 QAELENAY 211

RESULT 3

ABB07210

ID ABB07210 standard; Protein; 211 AA.

XX ABB07210;

AC ABB07210;

XX 26-MAR-2002 (first entry)

DT Mouse cloaked-2 polypeptide sequence.

XX Cloaked-2; cysteine knot motif; nephrotropic; cardiant; immunomodulator;

XX hepatotropic; antiinflammatory; antithyroid; cytoprotective;

XX antianemic; hypotensive; antiarrhythmic; cytoprotective; muscular;

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

GenCore version 5.1.4 p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 28, 2003, 14:06:06 ; Search time 51.8 Seconds
(without alignments)
475.895 Million cell updates/sec

Title: US-09-867-274-4

Perfect score: 1032
Sequence: 1 OCMQAFRNDATVIFGLGEY.....KGRKPRFGAKANQALEENAY 185

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq 101002:*

1: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*

2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*

3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*

4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*

5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*

6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*

7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*

8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*

9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*

10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*

11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*

12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*

13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*

14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*

15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*

16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*

17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*

18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*

19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*

20: /SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*

21: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*

22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*

23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1020.5	98.9	188	ABB07208	Mouse cloaked-2 po
2	1020.5	98.9	211	AA196432	Murine TGF-beta bi
3	1020.5	98.9	211	ABB07210	Mouse cloaked-2 po
4	982.5	95.2	213	AA196433	Rat TGF-beta bindi
5	951.5	92.2	213	AA196431	Vervet TGF-beta bi
6	932.5	90.4	190	ABB07207	Human cloaked-2 po
7	932.5	90.4	213	AA196430	Human DAN/Cerberus
8	932.5	90.4	213	AA196429	Human TGF-beta bin
9	932.5	90.4	213	AA197589	Human secreted pro
10	932.5	90.4	213	ABG34061	Human Pro peptide

11	932.5	90.4	213	ABB07209	Human cloaked-2 po
12	932.5	90.4	213	AA197089	Human osteolevin p
13	923.5	89.5	213	AA196430	Human TGF-beta bin
14	923.5	89.5	213	AA196436	Human TGF-beta bi
15	839	81.3	176	AA196434	Bovine TGF-beta bi
16	837.5	81.2	367	AA196435	Human DAN/Cerberus
17	687.5	66.6	139	AA196436	Human bone marrow
18	687.5	66.6	139	AA196436	Human bone marrow
19	330	32.0	206	AA196436	Human peptide enco
20	330	32.0	206	AA196436	Murine skin cell p
21	330	32.0	206	AA196436	Murine skin cell p
22	330	32.0	206	AA196436	Skin cell protein,
23	330	32.0	206	AA196436	Skin cell protein,
24	330	32.0	206	AA196436	Murine protein iso
25	330	32.0	206	AA196436	Murine protein iso
26	328	31.8	206	AA196436	Human small CCN-ii
27	328	31.8	206	AA196436	Human small CCN-ii
28	328	31.8	206	AA196436	Homo sapiens fetal
29	328	31.8	206	AA196436	Human adult retina
30	328	31.8	206	AA196436	Human angiogenesis
31	328	31.8	206	AA196436	Human small CCN-ii
32	328	31.8	206	AA196436	SCGF CNN family pr
33	324	31.4	183	AA196436	Human PRO532 prote
34	324	31.4	183	AA196436	Human cloaked-1 po
35	324	31.4	183	AA196436	Human breast tumou
36	324	31.4	183	AA196436	Human breast tumou
37	109.5	10.6	116	AA196436	Human 5' EST seque
38	109.5	10.6	116	AA196436	Secreted protein A
39	98.5	9.5	102	AA196436	Amino acid sequenc
40	98.5	9.5	102	AA196436	Human fetal kidney
41	90.5	8.8	184	AA196436	DAN and b57 protei
42	90.5	8.8	184	AA196436	Human polypeptide
43	90.5	8.8	184	AA196436	Human ORFX ORF2911
44	90.5	8.8	184	AA196436	Human polypeptide
45	88.5	8.6	182	AA196436	DAN and b57 protei

ALIGNMENTS

RESULT 1

ABB07208

ID ABB07208 standard; Protein; 188 AA.

XX

AC ABB07208;

XX

DT 26-MAR-2002 (first entry)

XX

DE Mouse cloaked-2 polypeptide mature protein sequence.

XX

DE Cloaked-2; cysteine knot motif; nephrotropic; cardiant; immunomodulator;

XX

KW hepatotropic; antiinflammatory; antithyroid; cytosolic; neuroprotective;

XX

KW antianemic; hypotensive; antiarrhythmic; antiarteriosclerotic; muscular;

XX

KW antididiabetic; anorectic; gene therapy; cell therapy; antisense therapy;

XX

XX mouse.

XX

OS Mus musculus.

XX

FT Key Location/Qualifiers

FT Misc-difference 175..177

FT /note= "this region is missing in the sequence provided

FT in the sequence listing but has been indicated

FT correctly in the sequence in the Figure"

XX

XX WO200192308-A2.

XX

XX 06-DEC-2001.

XX

XX 29-MAY-2001; 2001WO-US17478.

XX

XX 01-JUN-2000; 2000US-208550P.

XX

XX 04-AUG-2000; 2000US-223542F.

XX

; TITLE OF INVENTION: Interferon-Suppressing Placental Lactogen Peptides
; FILE REFERENCE: 1101

; CURRENT APPLICATION NUMBER: US/09/876,478
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: 60/210,082
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: MS-DOS
; SEQ ID NO 16
; LENGTH: 56
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
; FEATURE:
; OTHER INFORMATION: N-TERMINAL 56 RESIDUES OF HPRL
US-09-876-478-16

Query Match 3.2%; Score 6; DB 10; Length 56;
Best Local Similarity 100.0%; Pred. No. 72;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 123 CPGGAA 128
Db 32 CPGGAA 37
|||||

RESULT 14
US-09-510-332-72
; Sequence 72, Application US/09510332
; Publication No. US2003002278A1
; GENERAL INFORMATION:
; APPLICANT: Zuker, Charles S.
; APPLICANT: Adler, Jon Elliot
; APPLICANT: Ryba, Nick
; APPLICANT: Mueller, Ken
; APPLICANT: Hoon, Mark
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: T2R, a No. US2003002278A1el Family of Taste Receptors
; FILE REFERENCE: 02307E-098010US
; CURRENT APPLICATION NUMBER: US/09/510,332
; CURRENT FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 09/393,634
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 172
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 72
; LENGTH: 59
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human T2R46 (hGR46)
US-09-510-332-72

Query Match 3.2%; Score 6; DB 9; Length 59;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 77 ELVCSG 82
Db 54 ELVCSG 59
|||||

RESULT 15
US-09-864-761-35511
; Sequence 35511, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1

; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 35511
; LENGTH: 60
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AJ239318.3
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 3.3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.7
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.96
; OTHER INFORMATION: EST HUMAN HIT: AW962016.1, EVALUE 2.00e-20
; OTHER INFORMATION: SWISSPROT HIT: P41997, EVALUE 3.00e-05
US-09-864-761-35511

Query Match 3.2%; Score 6; DB 10; Length 60;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 157 FGPETA 162
Db 20 FGPETA 25
|||||

Search completed: March 28, 2003, 14:24:58
Job time : 12.84 secs

us-09-867-274-4.oligo.rapb

Mon Mar 31 09:44:33 2003

PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 34998
LENGTH: 41
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC007898.1
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.7
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.5
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.9
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.4
OTHER INFORMATION: EXPRESSED IN HELLO, SIGNAL = 2
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.2
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2
NAME/KEY: unsure
LOCATION: 4
NAME/KEY: unsure
LOCATION: 39
US-09-864-761-34998
Query Match 3.2%; Score 6; DB 10; Length 41;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 159 PETARP 164
Db 5 PETARP 10
RESULT 11
US-09-925-297-836
Sequence 836, Application US/09925297
Patent No. US20020081659A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PAL05
CURRENT APPLICATION NUMBER: US/09/925,297
CURRENT FILING DATE: 2001-08-10

PRIOR APPLICATION NUMBER: PCT/US00/05989
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 928
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 836
LENGTH: 46
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (4)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (6)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (28)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (42)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-297-836
Query Match 3.2%; Score 6; DB 10; Length 46;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 73 KPVTTEL 78
Db 34 KPVTTEL 39
RESULT 12
US-10-001-857-164
Sequence 164, Application US/10001857
Publication No. US20020183500A1
GENERAL INFORMATION:
APPLICANT: Macina, Roberto
APPLICANT: Recipon, Hervé
APPLICANT: Chen, Sei-Yu
APPLICANT: Sun, Yongming
APPLICANT: Liu, Chenghua
TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Proteins
FILE REFERENCE: DEX-0273
CURRENT APPLICATION NUMBER: US/10/001,857
CURRENT FILING DATE: 2001-11-20
PRIOR APPLICATION NUMBER: 60/252,054
PRIOR FILING DATE: 2000-11-20
NUMBER OF SEQ ID NOS: 208
SOFTWARE: PatentIn version 3.1
SEQ ID NO 164
LENGTH: 53
TYPE: PRT
ORGANISM: Homo sapiens
US-10-001-857-164
Query Match 3.2%; Score 6; DB 9; Length 53;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 37 GGRPPH 42
Db 21 GGRPPH 26
RESULT 13
US-09-876-478-16
Sequence 16, Application US/09876478
Patent No. US20020032154A1
GENERAL INFORMATION:
APPLICANT: Peyman, John A.

;; PRIOR APPLICATION NUMBER: PCT/US01/00667
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00664
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00669
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00665
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00668
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00663
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00662
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00661
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00670
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: US 09/774,203
;; PRIOR FILING DATE: 2001-01-29
;; NUMBER OF SEQ ID NOS: 49117
;; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
;; SEQ ID NO 44917
;; LENGTH: 26
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AC021000.3
;; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.76
;; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.95
;; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.86
;; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.93
US-09-864-761-44917

Query Match 3.2%; Score 6; DB 10; Length 26;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 171 RPKAKA 176
Db 7 RPKAKA 12

RESULT 9
US-09-864-761-40983
; Sequence 40983, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Acomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30

;; PRIOR APPLICATION NUMBER: PCT/US01/00667
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00664
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00669
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00665
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00668
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00663
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00662
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00661
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00670
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: US 09/774,203
;; PRIOR FILING DATE: 2001-01-29
;; NUMBER OF SEQ ID NOS: 49117
;; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
;; SEQ ID NO 40983
;; LENGTH: 40
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AC018742.2
;; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.8
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.7
;; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.9
;; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.3
;; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.3
;; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.8
;; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.5
;; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.2
;; OTHER INFORMATION: SWISSPROT HIT: P22297, EVALUE 2.20e+00
;; OTHER INFORMATION: EST_HUMAN HIT: A1220244.1, EVALUE 6.00e-03
US-09-864-761-40983

Query Match 3.2%; Score 6; DB 10; Length 40;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 25 PENNOT 30
Db 24 PENNOT 29

RESULT 10
US-09-864-761-34998
; Sequence 34998, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Acomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03

us-09-867-274-4.oligo.rapb

Mon Mar 31 09:44:33 2003

```

QY 65 TDGPCRSAPVTVELVCSGCGPARLLPNAIGR 96
| | | | | | | | | | | | | | | | | | | | |
Db 67 TDGPCRSAPVTVELVCSGCGPARLLPNAIGR 98
| | | | | | | | | | | | | | | | | | | | |

RESULT 5
US-09-867-274-5
; Sequence 5, Application US/09867274
; Patent No. US20020106650A1
; GENERAL INFORMATION:
; APPLICANT: Paszty, Christopher
; TITLE OF INVENTION: Cysteine Knot Polypeptides: Cloaked-2 Molecules and Uses Thereof
; FILE REFERENCE: 01017737428
; CURRENT APPLICATION NUMBER: US/09/867,274
; PRIOR FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: US 60/208,550
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: US 60/223,542
; PRIOR FILING DATE: 2000-08-04
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 5
; LENGTH: 213
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-867-274-5

Query Match 17.3%; Score 32; DB 10; Length 213;
Best Local Similarity 100.0%; Pred. No. 2.4e-24;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 65 TDGPCRSAPVTVELVCSGCGPARLLPNAIGR 96
| | | | | | | | | | | | | | | | | | | | |
Db 90 TDGPCRSAPVTVELVCSGCGPARLLPNAIGR 121
| | | | | | | | | | | | | | | | | | | | |

RESULT 6
US-09-732-680A-2
; Sequence 2, Application US/09732680A
; Patent No. US2002004612A1
; GENERAL INFORMATION:
; APPLICANT: Bayer Aktiengesellschaft
; TITLE OF INVENTION: Nucleic acids coding for new acetylcholine receptor beta subunits
; TITLE OF INVENTION: insects
; FILE REFERENCE: Le A 34 147
; CURRENT APPLICATION NUMBER: US/09/732,680A
; CURRENT FILING DATE: 2001-05-10
; PRIOR APPLICATION NUMBER: DE 199 59 582.8
; PRIOR FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 441
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-732-680A-2

Query Match 3.8%; Score 7; DB 10; Length 441;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 PGAKANQ 178
| | | | |
Db 404 PGAKANQ 410
| | | | |

RESULT 7
US-09-815-242-13801
; Sequence 13801, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert

QY 11 TEVIPGL 17
| | | | |
Db 72 TEVIPGL 78
| | | | |

Query Match 3.8%; Score 7; DB 10; Length 535;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 TEVIPGL 17
| | | | |
Db 72 TEVIPGL 78
| | | | |

RESULT 8
US-09-864-761-44917
; Sequence 44917, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30

```



```

; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annotmax Sequence Listing Engine vers. 1.1
; SEQ ID NO 47109
; LENGTH: 139
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC003098.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.62
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.6
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.49
; OTHER INFORMATION: SWISSPROT HIT: P45646, EVALUE 4.70e-01
US-09-867-761-47109

Query Match 17.3%; Score 32; DB 10; Length 139;
Best Local Similarity 100.0%; Pred. No. 1.6e-24;
Matches 32; Conservative 0; Mismatches 0; Indels

QY 65 TDGPCSAKPVTELVCISGCGPARLLPNAIGR 96
|||||
Db 16 TDGPCSAKPVTELVCISGCGPARLLPNAIGR 47

RESULT 4
US-09-867-274-2
; Sequence 2, Application US/09867274
; Patent No. US20020106650A1
; GENERAL INFORMATION:
; APPLICANT: Paszty, Christopher
; APPLICANT: Gao, Yongming
; TITLE OF INVENTION: Cysteine Knot Polypeptides: Cloaked-2 Mol
; FILE REFERENCE: 01017/3428
; CURRENT APPLICATION NUMBER: US/09/867,274
; CURRENT FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: US 60/208,550
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: US 60/223,542
; PRIOR FILING DATE: 2000-08-04
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 190
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-867-274-2

Query Match 17.3%; Score 32; DB 10; Length 190;
Best Local Similarity 100.0%; Pred. No. 2.1e-24;
Matches 32; Conservative 0; Mismatches 0; Indels

```

us-09-867-274-4.oligo.rapb

Mon Mar 31 09:44:33 2003

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 28, 2003, 14:22:02 ; Search time 11.84 seconds

(without alignments)
917.557 Million cell updates/sec

Title: US-09-867-274-4
Perfect score: 185
Sequence: 1 QGQAFRNDATVIFGLGEY.....KGRKPRFGAKANQAELENAY 185

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 237916 seqs, 58723674 residues

Word size : 0

Total number of hits satisfying chosen parameters: 237916

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications AA.*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB_PEP.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB_PEP.*
3: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB_PEP.*
4: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB_PEP.*
5: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB_PEP.*
6: /cgn2_6/ptodata/1/pubpaa/PCT05_PUBCOMB_PEP.*
7: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB_PEP.*
8: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB_PEP.*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB_PEP.*
10: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB_PEP.*
11: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB_PEP.*
12: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB_PEP.*
13: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB_PEP.*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB_PEP.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	185	100.0	185	US-09-867-274-4	Sequence 4, Appli
2	185	100.0	208	US-09-867-274-6	Sequence 6, Appli
3	32	17.3	139	US-09-864-761-47109	Sequence 47109, A
4	32	17.3	190	US-09-867-274-2	Sequence 2, Appli
5	32	17.3	213	US-09-867-274-5	Sequence 5, Appli
6	7	3.8	441	US-09-732-680A-2	Sequence 2, Appli
7	7	3.8	535	US-09-815-242-13801	Sequence 13801, A
8	6	3.2	26	US-09-864-761-44917	Sequence 44917, A
9	6	3.2	40	US-09-864-761-40983	Sequence 40983, A
10	6	3.2	41	US-09-864-761-34998	Sequence 34998, A
11	6	3.2	46	US-09-925-297-836	Sequence 836, App
12	6	3.2	53	US-10-001-857-164	Sequence 164, App
13	6	3.2	56	US-09-876-478-16	Sequence 72, Appli
14	6	3.2	59	US-09-510-332-72	Sequence 3511, A
15	6	3.2	70	US-09-864-761-35511	Sequence 878, App
16	6	3.2	74	US-09-925-297-878	Sequence 17, Appli
17	6	3.2	83	US-09-790-264-17	Sequence 47343, A
18	6	3.2	91	US-09-864-761-47343	Sequence 320, App
19	6	3.2	110	US-09-741-669-320	

20	6	3.2	110	US-09-912-020-387	Sequence 387, App
21	6	3.2	110	US-09-815-242-10354	Sequence 10354, A
22	6	3.2	110	US-09-815-242-13970	Sequence 13970, A
23	6	3.2	112	US-09-815-242-5235	Sequence 5235, Ap
24	6	3.2	113	US-10-013-379-17	Sequence 17, Appli
25	6	3.2	117	US-09-815-242-12258	Sequence 12258, A
26	6	3.2	117	US-09-815-242-12765	Sequence 12765, A
27	6	3.2	125	US-10-036-869-25	Sequence 25, Appli
28	6	3.2	137	US-09-867-550-832	Sequence 832, App
29	6	3.2	138	US-09-925-297-879	Sequence 879, App
30	6	3.2	139	US-09-764-870-521	Sequence 521, App
31	6	3.2	143	US-09-815-242-11669	Sequence 11669, A
32	6	3.2	146	US-09-764-877-1316	Sequence 1316, Ap
33	6	3.2	160	US-09-815-242-11664	Sequence 11664, A
34	6	3.2	163	US-10-002-769-8	Sequence 8, Appli
35	6	3.2	181	US-10-101-464A-587	Sequence 587, App
36	6	3.2	186	US-09-944-160-16	Sequence 16, Appli
37	6	3.2	191	US-09-764-870-417	Sequence 417, App
38	6	3.2	195	US-09-858-664A-11	Sequence 11, Appli
39	6	3.2	199	US-10-140-293-12	Sequence 12, Appli
40	6	3.2	203	US-09-815-242-5050	Sequence 5050, Ap
41	6	3.2	205	US-09-905-291A-23	Sequence 23, Appli
42	6	3.2	205	US-10-066-500-99	Sequence 99, Appli
43	6	3.2	205	US-09-902-853-23	Sequence 23, Appli
44	6	3.2	205	US-09-907-824-23	Sequence 23, Appli
45	6	3.2	205	US-09-907-841-23	Sequence 23, Appli

ALIGNMENTS

RESULT 1
US-09-867-274-4
; Sequence 4, Application US/09867274
; Patent No. US20020106650A1
; GENERAL INFORMATION:
; APPLICANT: Paszty, Christopher
; APPLICANT: Gao, Yongming
; TITLE OF INVENTION: Cysteine Knot Polypeptides: Cloaked-2 Molecules and Uses Thereof
; FILE REFERENCE: 01017/37428
; CURRENT APPLICATION NUMBER: US/09/867,274
; CURRENT FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: US 60/208,550
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: US 60/223,542
; PRIOR FILING DATE: 2000-08-04
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 185
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-867-274-4

Query Match	100.0%	Score 185;	DB 10;	Length 185;
Best Local Similarity	100.0%	Pred. No. 1.8e-177;		
Matches 185;	Conservative	0;	Mismatches 0;	Indels 0; Gaps 0;
QY	1	QGQAFRNDATVIFGLGEYPEPPENNQTWNRAENGRRPHHPYDAKDVSSEYSCRELHY	60	
Db	1	QGQAFRNDATVIFGLGEYPEPPENNQTWNRAENGRRPHHPYDAKDVSSEYSCRELHY	60	
QY	61	TRFLTDPGCRSAKPVTELVCSCGCGPARLLPNAIGRVKWRPNPGDFPFCIPDRYARQVQ	120	
Db	61	TRFLTDPGCRSAKPVTELVCSCGCGPARLLPNAIGRVKWRPNPGDFPFCIPDRYARQVQ	120	
QY	121	LLCPGGAAPRRKRVLRVASCCKRLTRFHNSQLKDFGPETARPOKGRKPRFGAKANQAE	180	
Db	121	LLCPGGAAPRRKRVLRVASCCKRLTRFHNSQLKDFGPETARPOKGRKPRFGAKANQAE	180	
QY	181	LENAY 185		
Db	181	LENAY 185		

RESULT 15
 US-08-600-993A-30
 ; Sequence 30, Application US/08600993A
 ; Patent No. 5981229
 ; GENERAL INFORMATION:
 ; APPLICANT: Masure, H Robert
 ; APPLICANT: Pearce, Barbara J
 ; APPLICANT: Tuomanen, Elaine
 ; TITLE OF INVENTION: BACTERIAL EXPORTED PROTEINS AND
 ; TITLE OF INVENTION: ACCELLULAR VACCINES BASED THEREON
 ; NUMBER OF SEQUENCES: 59
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Klauber & Jackson
 ; STREET: 411 Hackensack Avenue
 ; CITY: Hackensack
 ; STATE: New Jersey
 ; COUNTRY: USA
 ; ZIP: 07601
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/600,993A
 ; FILING DATE: 1-MAR-1996
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/245,511
 ; FILING DATE: 18-MAY-1994
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/116,541
 ; FILING DATE: 01-SEP-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Jackson Esq., David A.
 ; REGISTRATION NUMBER: 26,742
 ; REFERENCE/DOCKET NUMBER: 600-1-069 US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 201 487-5800
 ; TELEFAX: 201 343-1684
 ; TELEX: 133521
 ; INFORMATION FOR SEQ ID NO: 30:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 97 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: unknown
 ; MOLECULE TYPE: peptide
 ; HYPOTHETICAL: NO
 ; ANTI-SENSE: NO
 ; FRAGMENT TYPE: internal
 ; ORIGINAL SOURCE:
 ; ORGANISM: Bacillus subtilis
 ; US-08-600-993A-30

Query Match 3.2%; Score 6; DB 2; Length 97;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 167 GRKRP 172
 Db 80 GRKRP 85

Search completed: March 28, 2003, 14:24:27
 Job time : 11.3467 secs

us-09-867-274-4.oligo.ra1

Mon Mar 31 09:44:32 2003

```
;
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
US-08-836-791-5
Query Match 3.2%; Score 6; DB 2; Length 68;
Best Local Similarity 100.0%; Pred. No. 90;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 152 SELKDF 157
DB 35 SELKDF 40
RESULT 14
US-08-245-511-30
; Sequence 30, Application US/08245511
; Patent No. 5928900
; GENERAL INFORMATION:
; APPLICANT: Masure, H Robert
; APPLICANT: Pearce, Barbara J
; APPLICANT: Tuomanen, Elaine
; TITLE OF INVENTION: BACTERIAL EXPORTED PROTEINS AND
; TITLE OF INVENTION: ACCELLULAR VACCINES BASED THEREON
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM: disk
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/245,511
; FILING DATE: 18-MAY-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/116,541
; FILING DATE: 01-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-069 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 97 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; ORGANISM: Bacillus subtilis
US-08-245-511-30
Query Match 3.2%; Score 6; DB 2; Length 97;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 167 GRKPRP 172
DB 80 GRKPRP 85
;
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: HU91-06A2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 24 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-220-272A-6
Query Match 3.2%; Score 6; DB 1; Length 24;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 127 AAPRSR 132
DB 13 AAPRSR 18
RESULT 13
US-08-836-791-5
; Sequence 5, Application US/08836791
; Patent No. 5886150
; GENERAL INFORMATION:
; APPLICANT: Duchesne, Marc
; APPLICANT: Faucher, Didier
; APPLICANT: Parker, Fabienne
; APPLICANT: Schweighoffer, Fabien
; APPLICANT: Tocque, Bruno
; TITLE OF INVENTION: Peptides Capable of Binding to the GAP
; TITLE OF INVENTION: Protein SH3 Domain, Nucleotide Sequences Coding Therefor,
; TITLE OF INVENTION: and Preparation and Use Thereof
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, Mailstop 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/836,791
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 94/13955
; FILING DATE: 22-NOV-1994
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 68 amino acids
; TYPE: amino acid
; NAME: Fehner Esq., Paul F.
; REGISTRATION NUMBER: 35,135
; REFERENCE/DOCKET NUMBER: ST94083G1-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3839
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 68 amino acids
; TYPE: amino acid
```

Query Match 3.2%; Score 6; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 PGLGEY 20
Db 4 PGLGEY 9

RESULT 10

US-08-250-789A-96
; Sequence 96, Application US/08250789A
; Patent No. 5635597
; GENERAL INFORMATION:
; APPLICANT: Barrett, Ronald W.
; APPLICANT: Chernov-Rogan, Tania
; APPLICANT: Davis, Ann M.
; TITLE OF INVENTION: Peptides That Bind to IL-2 Receptors
; NUMBER OF SEQUENCES: 194
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: One Market Plaza, Steuart Tower, Suite 2000
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94105

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/250,789A
; FILING DATE: 27-MAY-1994
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5635597v1el, Vernon A.
; REGISTRATION NUMBER: 32,483
; REFERENCE/DOCKET NUMBER: 16528A-57/1043
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-326-2400
; TELEFAX: 415-326-2422
; INFORMATION FOR SEQ ID NO: 96:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-250-789A-96

Query Match 3.2%; Score 6; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 PGLGEY 20
Db 4 PGLGEY 9

RESULT 11

US-08-250-789A-97
; Sequence 97, Application US/08250789A
; Patent No. 5635597
; GENERAL INFORMATION:
; APPLICANT: Barrett, Ronald W.
; APPLICANT: Chernov-Rogan, Tania
; APPLICANT: Davis, Ann M.
; TITLE OF INVENTION: Peptides That Bind to IL-2 Receptors
; NUMBER OF SEQUENCES: 194
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew

; STREET: One Market Plaza, Steuart Tower, Suite 2000
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94105

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/250,789A
; FILING DATE: 27-MAY-1994
; CLASSIFICATION: 514

; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5635597v1el, Vernon A.
; REGISTRATION NUMBER: 32,483
; REFERENCE/DOCKET NUMBER: 16528A-57/1043
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-326-2400
; TELEFAX: 415-326-2422
; INFORMATION FOR SEQ ID NO: 97:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-250-789A-97

Query Match 3.2%; Score 6; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 PGLGEY 20
Db 4 PGLGEY 9

RESULT 12

US-08-220-272A-6
; Sequence 6, Application US/08220272A
; Patent No. 5783384
; GENERAL INFORMATION:
; APPLICANT: Verdine, Gregory L.
; TITLE OF INVENTION: Selection of Binding Molecule
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: US
; ZIP: 02173

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/220,272A
; FILING DATE: 30-MAR-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/029,331
; FILING DATE: February 10, 1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/819,855
; FILING DATE: January 13, 1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/00321
; FILING DATE: January 13, 1993
; ATTORNEY/AGENT INFORMATION:

```

; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 213
; TYPE: PRT
; ORGANISM: Homo sapien
; US-09-449-218D-8

Query Match          17.3%; Score 32; DB 4; Length 213;
Best Local Similarity 100.0%; Pred. No. 3.8e-24;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      65  TDGPCRSAPVTVELVCSGGCGPARLLPNAIGR 96
|||||
DB      90  TDGPCRSAPVTVELVCSGGCGPARLLPNAIGR 121
|||||

```

US-09-449-218D-10
; Sequence 10, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.

```

; APPLICANT:  van Nessel, David G.
; APPLICANT:  Winkler, David G.
; TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR INCREASING
; TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR INCREASING
; TITLE OF INVENTION:  BONE MINERALIZATION
; TITLE OF INVENTION:  BONE MINERALIZATION
; FILE REFERENCE:  240083.508
; CURRENT APPLICATION NUMBER:  US/09/449,218D
; CURRENT FILING DATE:  1999-11-24
; NUMBER OF SEQ ID NOS:  45
; SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 10
; LENGTH:  213
; TYPE:  PRT
; ORGANISM:  Cercopithecus pygerythrus
US-09-449-218D-10

Query Match          17.3%;  Score 32;  DB 4;  Length 213;
Best Local Similarity 100.0%;  Pred. No. 3.8e-24;
                    100.0%; 0.0;  Mismatches 0;  Indels 0;  Gaps 0;

```

Matches 32; Conservative 0; Intermediate 0

QY 65 TDGPCRS AKPVTEL VCSGCG PARLLPNAIGR 96
|||||
Db 90 TDGPCRS AKPVTEL VCSGCG PARLLPNAIGR 121
|||||

RESULT 8
US-08-250-789A-94
; Sequence 94, Application US/08250789A
; Patent No. 5635597
; GENERAL INFORMATION:
; APPLICANT: Barrett, Ronald W.
; APPLICANT: Chernov-Rogan, Tania
; APPLICANT: Davis, Ann M.
; TITLE OF INVENTION: Peptides That Bind to IL-2 Receptors
; NUMBER OF SEQUENCES: 194
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Townsend and Townsend Khourie and Crew
; STREET: One Market Plaza, Steuart Tower, Suite 2000
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94105
; COMPILED READABLE FORM:

TIS-08-250-789A-95

```

; APPLICANT: Mulligan, John T.
; APPLICANT: Paepker, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 213
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-449-218D-14

```

```

Query Match      58.4%; Score 108; DB 4; Length 213;
Best Local Similarity 100.0%; Pred. No. 1.4e-99;
Matches 108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 65 TDGCRSAKPVTELVCSGCGPARLLPNAIGRVKWRPNPDPFCIPDRYRAQRVOLLCP 124
Db 90 TDGCRSAKPVTELVCSGCGPARLLPNAIGRVKWRPNPDPFCIPDRYRAQRVOLLCP 149

Qy 125 GGAAPRSRKVRLVASCKKRLTRFHQSELKDFGPETARPQGRKPRP 172
Db 150 GGAAPRSRKVRLVASCKKRLTRFHQSELKDFGPETARPQGRKPRP 197

```

```

RESULT 3
US-09-449-218D-16
; Sequence 16, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepker, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 16
; LENGTH: 176
; TYPE: PRT
; ORGANISM: Bos taurus
US-09-449-218D-16

```

```

Query Match      17.3%; Score 32; DB 4; Length 176;
Best Local Similarity 100.0%; Pred. No. 3.2e-24;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 65 TDGCRSAKPVTELVCSGCGPARLLPNAIGR 96
Db 59 TDGCRSAKPVTELVCSGCGPARLLPNAIGR 90

```

```

RESULT 4
US-09-449-218D-2
; Sequence 2, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.

```

```

; APPLICANT: Paepker, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 213
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-449-218D-2

```

```

Query Match      17.3%; Score 32; DB 4; Length 213;
Best Local Similarity 100.0%; Pred. No. 3.8e-24;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 65 TDGCRSAKPVTELVCSGCGPARLLPNAIGR 96
Db 90 TDGCRSAKPVTELVCSGCGPARLLPNAIGR 121

```

```

RESULT 5
US-09-449-218D-6
; Sequence 6, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepker, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 213
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-449-218D-6

```

```

Query Match      17.3%; Score 32; DB 4; Length 213;
Best Local Similarity 100.0%; Pred. No. 3.8e-24;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 65 TDGCRSAKPVTELVCSGCGPARLLPNAIGR 96
Db 90 TDGCRSAKPVTELVCSGCGPARLLPNAIGR 121

```

```

RESULT 6
US-09-449-218D-8
; Sequence 8, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepker, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 213
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-449-218D-8

```

us-09-867-274-4.oligo.ra1

Mon Mar 31 09:44:32 2003

GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 28, 2003, 14:20:52 ; Search time 11.3467 Seconds
(without alignments)
479.722 Million cell updates/sec

Title: US-09-867-274-4
Perfect score: 185
Sequence: 1 CQWQAFRNDATVIGLGEY.....KGRKPRGAKANQALENAY 185

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 262574 seqs, 29422922 residues

Word size : 0

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/1/iaa/5A COMB.pcp.*
2: /cgn2_6/ptodata/1/iaa/5B COMB.pcp.*
3: /cgn2_6/ptodata/1/iaa/6A COMB.pcp.*
4: /cgn2_6/ptodata/1/iaa/6B COMB.pcp.*
5: /cgn2_6/ptodata/1/iaa/PTCUTS COMB.pcp.*
6: /cgn2_6/ptodata/1/iaa/backfiles.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	174	94.1	211	4	US-09-449-218D-12
2	108	58.4	213	4	US-09-449-218D-14
3	32	17.3	176	4	US-09-449-218D-16
4	32	17.3	213	4	US-09-449-218D-2
5	32	17.3	213	4	US-09-449-218D-6
6	32	17.3	213	4	US-09-449-218D-8
7	32	17.3	213	4	US-09-449-218D-10
8	6	3.2	10	1	US-08-250-789A-94
9	6	3.2	10	1	US-08-250-789A-95
10	6	3.2	10	1	US-08-250-789A-96
11	6	3.2	10	1	US-08-250-789A-97
12	6	3.2	24	1	US-08-220-272A-6
13	6	3.2	68	2	US-08-836-791-5
14	6	3.2	97	2	US-08-245-511-30
15	6	3.2	97	2	US-08-600-993A-30
16	6	3.2	123	4	US-09-562-737-126
17	6	3.2	123	4	US-09-134-001C-3283
18	6	3.2	125	3	US-08-985-526-25
19	6	3.2	128	1	US-08-866-798-2
20	6	3.2	128	1	US-08-892-652-2
21	6	3.2	128	2	US-09-096-071-2
22	6	3.2	129	4	US-09-325-932A-50
23	6	3.2	182	4	US-09-230-670C-6
24	6	3.2	186	4	US-09-475-316A-13
25	6	3.2	187	4	US-09-134-001C-2866
26	6	3.2	199	3	US-08-737-248-7
27	6	3.2	199	4	US-09-325-932A-46

Sequence 127, Appl
Sequence 13, Appl
Sequence 13, Appl
Sequence 3806, Appl
Sequence 3, Appl
Sequence 10, Appl
Sequence 10, Appl
Sequence 27, Appl
Sequence 3304, Appl
Sequence 3, Appl
Sequence 2, Appl
Sequence 94, Appl
Sequence 144, Appl
Sequence 278, Appl
Sequence 13, Appl
Sequence 1, Appl
Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-09-449-218D-12
; Sequence 12, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepker, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; TITLE OF INVENTION: BONE MINERALIZATION
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 12
; LENGTH: 211
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-449-218D-12

Query Match 94.1%; Score 174; DB 4; Length 211;
Best Local Similarity 100.0%; Pred. No. 4.5e-165; Indels 0; Gaps 0;
Matches 174; Conservative 0; Mismatches 0;
QY 1 CQWQAFRNDATVIGLGEYPPPPENNQTWRAENGRRPPHPHYDAKOVSEYSCRELHY 60
DB 24 CQWQAFRNDATVIGLGEYPPPPENNQTWRAENGRRPPHPHYDAKOVSEYSCRELHY 83
QY 61 TRFLTDGCRSAKPYTELVCSCGQCPARLLPNAIGRVKWRPNRNGPFRICPDYRAQRVQ 120
DB 84 TRFLTDGCRSAKPYTELVCSCGQCPARLLPNAIGRVKWRPNRNGPFRICPDYRAQRVQ 143
QY 121 LLCPGAAPRSKRVLRVASCCKRLTRFHNSKLVKDFGPETARPKGRKPRPGA 174
DB 144 LLCPGAAPRSKRVLRVASCCKRLTRFHNSKLVKDFGPETARPKGRKPRPGA 197

RESULT 2
US-09-449-218D-14
; Sequence 14, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Marin E., Nussbaum L., Vavasseur A., Forestier C.;
 RT "POP: an Arabidopsis NBD-like encoding sequence.";
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=COLUMBIA;
 RX MEDLINE=97471969; PubMed=9330910;
 RA Sato S., Kotani H., Nakamura Y., Kaneko T., Asamizu E., Fukami M.,
 RA Miyajima N., Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. I. Sequence
 RT features of the 1.6 Mb regions covered by twenty physically assigned
 RT Pl clones.";
 RL DNA Res. 4:215-230(1997).
 DR EMBL; AF127664; AAD20643.1; -;
 DR EMBL; AB005239; BAB10978.1; -;
 DR InterPro; IPR003439; ABC_transportr.
 DR Pfam; PF00005; ABC_cran; 1.
 SQ SEQUENCE 282 AA; 31360 MW; 6F0E88F3EFDB1577 CRC64;

Query Match 3.8%; Score 7; DB 10; Length 282;
 Best Local Similarity 100.0%; Pred. No. 54;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 76 TELVCSG 82
 Db 82 TELVCSG 88

Search completed: March 28, 2003, 14:23:24
 Job time : 28.1467 secs

Mon Mar 31 09:44:34 2003

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 VIPGLGE 19
| | | | |
Db 222 VIPGLGE 228

RESULT 13

Q9L1NO PRELIMINARY; PRT; 251 AA.
AC Q9L1NO; 2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE Putative DNA methylase.
GN SCO6885 OR SC7F9.37.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Seeger K.J., Harris D.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Thomson N.R., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Redenbach M., Kieser H.M., Denapaita D., Eichner A., Cullum J.,
RL "A set of ordered cosmids and a detailed genetic and physical map for the 8 Mb Streptomyces coelicolor A3(2) chromosome."
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Batenan A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.H., Kieser T., Larke L., Murphy L., Oliver K., O'Neil S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces coelicolor A3(2).";
RL Nature 417:141-147(2002).
DR EMBL; AL138977; CAB22330.1; -.
DR InterPro; IPR001091; C24 Mettransf.
DR InterPro; IPR002295; D21N6_Mtase.
DR InterPro; IPR002941; N6/N4_Mtase.
DR InterPro; IPR002052; N6_Mtase.
DR Pfam; PF01555; N6_N4_Mtase; 1.
DR PRINTS; PR00506; D21N6MTFRASE.
DR PRINTS; PR00508; S21N4MTFRASE.
DR PROSITE; PS00092; N6_MTASE; UNKNOWN_1.
KW Methyltransferase.
SQ SEQUENCE 251 AA; 27405 MW; 1AEC919125AF9CE9 CRC64;

Query Match 3.8%; Score 7; DB 16; Length 251;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 162 ARPOKGR 168
| | | | |
Db 127 ARPOKGR 133

RESULT 14
Q9KZ52 PRELIMINARY; PRT; 278 AA.
AC Q9KZ52;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE Putative insertion element transposase.
GN SCO7080 OR SC3A4.06C.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Brown S.P., Harris D.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Thomson N.R., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Redenbach M., Kieser H.M., Denapaita D., Eichner A., Cullum J.,
RL "A set of ordered cosmids and a detailed genetic and physical map for the 8 Mb Streptomyces coelicolor A3(2) chromosome."
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Batenan A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.H., Kieser T., Larke L., Murphy L., Oliver K., O'Neil S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces coelicolor A3(2).";
RL Nature 417:141-147(2002).
DR EMBL; AL138977; CAB22330.1; -.
DR InterPro; IPR001091; C24 Mettransf.
DR InterPro; IPR002295; D21N6_Mtase.
DR InterPro; IPR002941; N6/N4_Mtase.
DR InterPro; IPR002052; N6_Mtase.
DR Pfam; PF01555; N6_N4_Mtase; 1.
DR PRINTS; PR00506; D21N6MTFRASE.
DR PRINTS; PR00508; S21N4MTFRASE.
DR PROSITE; PS00092; N6_MTASE; UNKNOWN_1.
KW Methyltransferase.
SQ SEQUENCE 278 AA; 30918 MW; 81564528BF1A2E12 CRC64;

Query Match 3.8%; Score 7; DB 16; Length 278;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 168 RKPRPGA 174
| | | | |
Db 209 RKPRPGA 215

RESULT 15
Q9XF19 PRELIMINARY; PRT; 282 AA.
AC Q9XF19;
DT 01-NOV-1999 (TRENBLrel. 12, Created)
DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE NBD-like protein.
GN POP.
OS Arabidopsis thaliana (Mouse-ear cress).

GN NM1656.
 OS Neisseria meningitidis (serogroup B).
 OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
 OX NCBI_TaxID=491;
 RN [1]
 RC SEQUENCE FROM N.A.
 RX STRAIN=MC58 / SEROGROUP B;
 MEDLINE=20175755; PubMed=10710307;
 RA Tetzelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
 Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
 Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
 Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
 Mason T., Ciecko A., Parksey D.S., Blair E., Ciftone H., Clark E.B.,
 Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J.,
 Gill J., Scarlato V., Masiagnani V., Pizzo M., Grandi G., Sun L.,
 Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
 RT "Complete genome sequence of Neisseria meningitidis serogroup B strain
 MC58.";
 RL Science 287:1809-1815(2000).
 DR EMBL; AE002516; AAF42005.1; -.
 DR TIGR; NM1656; -.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 197 AA; 21804 MW; B6C59F55DEAESAC8 CRC64;

Query Match 3.8%; Score 7; DB 16; Length 197;
 Best Local Similarity 100.0%; Pred. No. 40;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 175 KANQAE181
 Db 122 KANQAE128

RESULT 10

Q9JTA0
 ID Q9JTA0 PRELIMINARY; PRT; 197 AA.
 AC Q9JTA0
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Hypothetical protein NMA1913.
 GN NMA1913.
 OS Neisseria meningitidis (serogroup A).
 OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
 OX NCBI_TaxID=65699;
 RN [1]
 RC SEQUENCE FROM N.A.
 RX STRAIN=Z2491 / SEROGROUP A / SEROTYPE 4A;
 MEDLINE=20222556; PubMed=10761919;
 RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
 Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
 Davies R.W., Davis P., Devlin K., Feltwell T., Hamlin N., Hoiroyd S.,
 Jagels K., Leather S., Moule S., Mungall K., Quail M.A.,
 Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
 Whitehead S., Spratt B.G., Barrall B.G.;
 RT "Complete DNA sequence of a serogroup A strain of Neisseria
 meningitidis Z2491.";
 RL Nature 404:502-506(2000).
 DR EMBL; AL162757; CAB85134.1; -.
 DR Hypothetical protein; Complete proteome.
 SQ SEQUENCE 197 AA; 21747 MW; B6C5977D4CE35AC8 CRC64;

Query Match 3.8%; Score 7; DB 16; Length 197;
 Best Local Similarity 100.0%; Pred. No. 40;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 175 KANQAE181
 Db 122 KANQAE128

RESULT 11

Q9IBJ1
 ID Q9IBJ1 PRELIMINARY; PRT; 239 AA.
 AC Q9IBJ1
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE R-LORF1.
 GN R-LORF1.
 OS Turkey herpesvirus.
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Marek's disease-like viruses.
 OX NCBI_TaxID=10390;
 RN [1]
 RC SEQUENCE FROM N.A.
 RX STRAIN=GA;
 MEDLINE=92237304; PubMed=1315048;
 RA Jones D., Lee L., Liu J.L., Kung H.J., Tillotson J.K.;
 RT "Marek disease virus encodes a basic-leucine zipper gene resembling
 the fos/jun oncogenes that is highly expressed in lymphoblastoid
 tumors.";
 RL Proc. Natl. Acad. Sci. U.S.A. 89:4042-4046(1992).
 RN [2]
 RC SEQUENCE FROM N.A.
 RX STRAIN=GA;
 RA Lee L.F., Wu P., Sui D., Ren D., Kung H.J.;
 RT "The Complete UL Sequence of Serotype I Marek's Disease Virus.";
 RL Proc. Natl. Acad. Sci. U.S.A. 0:0-0(2000).
 DR EMBL; AF147806; AAF67204.1; -.
 DR EMBL; AF147806; AAF6792.1; -.
 DR InterPro: IPR000564; 2Fe2S; ferredoxin.
 DR PROSITE: PS00197; 2Fe2S-FERREDOXIN; UNKNOWN 1.
 SQ SEQUENCE 239 AA; 24655 MW; 922B2C2B2950F01D CRC64;

Query Match 3.8%; Score 7; DB 12; Length 239;
 Best Local Similarity 100.0%; Pred. No. 47;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 123 CPGGAAP 129
 Db 139 CPGGAAP 145

RESULT 12

Q8RXX1
 ID Q8RXX1 PRELIMINARY; PRT; 241 AA.
 AC Q8RXX1
 DT 01-JUN-2002 (TREMBlrel. 21, Created)
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Putative uridine kinase (Fragment).
 GN AT5G40870.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RC SEQUENCE FROM N.A.
 RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
 Goldsmith A.D., Lee J.M., Quach H.L., Toriumi M., Yu G., Bowser L.,
 Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
 Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
 Miranda M., Narusaka N., Nguyen M., Palm C.J., Sakurai T., Satou M.,
 Seki M., Shimizu P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,
 Theologis A.;
 RT "Arabidopsis Full Length cDNA Clones.";
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY080631; AAL85977.1; -.
 KW Kinase.
 FT NON TER
 SQ SEQUENCE 241 AA; 27003 MW; 266CF01112CEBEDS CRC64;

Query Match 3.8%; Score 7; DB 10; Length 241;
 Best Local Similarity 100.0%; Pred. No. 48;

Mon Mar 31 09:44:34 2003

|||||
Db 6 PGAKANQ 12

RESULT 6

Q9HVX4 PRELIMINARY; PRT; 148 AA.
ID Q9HVX4
AC Q9HVX4
DT 01-MAR-2001 (T-EMBLrel. 16, Created)
DT 01-MAR-2001 (T-EMBLrel. 16, Last sequence update)
DT 01-OCT-2001 (T-EMBLrel. 18, Last annotation update)
DE Hypothetical protein PA4441.
GN PA4441.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OC NCBI_TaxID=287;
OX [1]
RN SEQUENCE FROM N.A.
RP STRAIN=ATCC 15692 / PA01;
RC MEDLINE=20437337; PubMed=10984043;
RX Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
Hickey M.J., Brinkman P.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RA "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen."
RL Nature 406:959-964 (2000).
DR EMBL; AE004858; AAG07829.1; -;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 148 AA; 16418 MW; 4DD9B4B217EDD2B8 CRC64;

Query Match 3.8%; Score 7; DB 16; Length 148;

Best Local Similarity 100.0%; Pred. No. 31; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0;

QY 87 ARLLPNA 93
|||||

Db 26 ARLLPNA 32

RESULT 7

Q8Z095 PRELIMINARY; PRT; 162 AA.
ID Q8Z095
AC Q8Z095
DT 01-MAR-2002 (T-EMBLrel. 20, Created)
DT 01-MAR-2002 (T-EMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (T-EMBLrel. 20, Last annotation update)
DE Hypothetical protein ALR0203.
GN ALR0203.
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OC NCBI_TaxID=103690;
OX [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=21595285; PubMed=11759840;
RX Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
Nakazaki N., Shimo S., Sugimoto M., Takazawa M., Yamada M.,
Yasuda M., Tabata S.;
RA "Complete genomic sequence of the filamentous nitrogen-fixing
RT cyanobacterium Anabaena sp. strain PCC 7120."
RL DNA Res. 8:205-213 (2001).
DR EMBL; AP003581; BAB7727.1; -;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 162 AA; 18240 MW; 7B56B0FDE5E663C79 CRC64;

Query Match 3.8%; Score 7; DB 16; Length 162;

Best Local Similarity 100.0%; Pred. No. 34; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0;

QY 172 PGAKANQ 178

RESULT 8

Q9AK55 PRELIMINARY; PRT; 169 AA.
ID Q9AK55
AC Q9AK55
DT 01-JUN-2001 (T-EMBLrel. 17, Created)
DT 01-JUN-2001 (T-EMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)
DE Putative transposase.
GN SC04062 OR 2SCD60.28.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OC NCBI_TaxID=1902;
OX [1]
RN SEQUENCE FROM N.A.
RP STRAIN=A3(2);
RC Harris D.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Cerdeno A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
RA Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RX MEDLINE=97000351; PubMed=8843436;
RA Redenbach M., Kieser H.M., Denapaite D., Eichner A., Cullum J.,
Kinaishi H., Hopwood D.A.;
RA "A set of ordered cosmids and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome."
RL Mol. Microbiol. 21:77-96 (1996).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neil S.,
Rabinowitz E., Rajandream M.A., Rutherford K., Rutter S., Taylor K.,
Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor J.,
Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
Hopwood D.A.;
RA "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2)."
RL Nature 417:141-147 (2002).
DR EMBL; AL583944; CAC32333.1; -;
DR InterPro; IPR002559; Transposase 11.
DR Pfam; PF01609; Transposase 11; 1.
SQ SEQUENCE 169 AA; 18628 MW; A69871DB6E1902FE CRC64;

Query Match 3.8%; Score 7; DB 16; Length 169;

Best Local Similarity 100.0%; Pred. No. 35; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0;

QY 39 RPPHPY 45
|||||

Db 153 RPPHPY 159

RESULT 9

Q9UYB9 PRELIMINARY; PRT; 197 AA.
ID Q9UYB9
AC Q9UYB9
DT 01-OCT-2000 (T-EMBLrel. 15, Created)
DT 01-OCT-2000 (T-EMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (T-EMBLrel. 19, Last annotation update)
DE Hypothetical protein NMB1656.

Qy	TEVIPGL 17
11	
62	TEVIPGL 68

Y	133	KVRLVAS	139
b	12	KVRLVAS	18

Mon Mar 31 09:44:34 2003

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 28, 2003, 14:19:52 ; Search time 26.1467 Seconds
(without alignments)
1457.881 Million cell updates/sec

Title: US-09-867-274-4
Perfect score: 185
Sequence: 1 QGWAQFNDATEVPLGLGV.....KGRKPRGAKANQAELENAY 185

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 671580 seqs, 206047115 residues

Word size : 0
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

- Database : SPTREMBL 21.*
- 1: sp_archaea.*
 - 2: sp_bacteria.*
 - 3: sp_fungi.*
 - 4: sp_human.*
 - 5: sp_invertebrate.*
 - 6: sp_mammal.*
 - 7: sp_mhc.*
 - 8: sp_organelle.*
 - 9: sp_phase.*
 - 10: sp_plant.*
 - 11: sp_rodent.*
 - 12: sp_virus.*
 - 13: sp_vertebrate.*
 - 14: sp_unclassified.*
 - 15: sp_rvirus.*
 - 16: sp_bacteriap.*
 - 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8	4.3	428	Q969X8	Q969X8 homo sapien
2	7	3.8	89	Q9FH55	Q9FH55 arabidopsis
3	7	3.8	89	Q8ZJ53	Q8ZJ53 yersinia pe
4	7	3.8	109	Q9GTJ4	Q9GTJ4 ancylostoma
5	7	3.8	130	Q910C0	Q910C0 pseudomonas
6	7	3.8	148	Q9HVX4	Q9HVX4 pseudomonas
7	7	3.8	162	Q8Z095	Q8Z095 anabaena sp
8	7	3.8	169	Q9AK55	Q9AK55 streptomyce
9	7	3.8	197	Q9JYB9	Q9JYB9 neisseria m
10	7	3.8	197	Q9JTA0	Q9JTA0 neisseria m
11	7	3.8	239	Q9IBJ1	Q9IBJ1 turkey herp
12	7	3.8	241	Q8RXX1	Q8RXX1 arabidopsis
13	7	3.8	251	Q9LJN0	Q9LJN0 streptomyce
14	7	3.8	278	Q9KZ52	Q9KZ52 streptomyce
15	7	3.8	282	Q9XF19	Q9XF19 arabidopsis
16	7	3.8	311	Q8UBG9	Q8UBG9 agrobacteri

Q9PQ72 ureaplasma
Q9FUN1 zea mays (m
Q915Y7 pseudomonas
Q91X3 streptomyce
Q91G89 oryza sativ
Q90388 methanococc
Q95K71 macaca fasc
Q19348 caenorhabdi
Q9VPQ8 drosophila
Q8S485 zea mays (m
Q91Y6 arabidopsis
Q91K34 arabidopsis
Q9F80 arabidopsis
Q9HVR9 pseudomonas
Q8ZLA9 salmonella
Q8Z299 salmonella
Q8X4K7 escherichia
Q9594 caenorhabdi
Q91837 arabidopsis
P95077 mycobacteri
Q44003 toxoplasma
Q92D38 listeria in
Q9H7Y7 homo sapien
Q85050 pseudorabie
Q9PGB8 xylella fas
Q9BSC4 homo sapien
Q8XPT2 raistonia s
Q9ABT1 caulobacter
Q90X95 xenopus lae

ALIGNMENTS

RESULT 1

Q969X8 PRELIMINARY; PRT; 428 AA.
ID Q969X8
AC Q969X8;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Elkl protein.
GN ELK1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=HIPPOCAMPUS;
RX MEDLINE=99246057; PubMed=10231026;
RA Yamauchi T., Toko M., Suga M., Hatakeyama T., Isobe M.;
RT "Structural organization of the human ELK1 gene and its processed
RT pseudogene ELK2 genes.";
RL DNA Res. 6:21-27(1999).
DR EMBL; AB016194; BAA36617.1; -;
DR EMBL; AB016193; BAA36616.1; -;
DR InterPro; IPR000418; Ets.
DR InterPro; IPR002341; HSP_ETs.
DR Pfam; PF00178; Ets; 1.
DR PROSITE; PS00345; ETS_DOMAIN_1; UNKNOWN_1.
DR PROSITE; PS00346; ETS_DOMAIN_2; UNKNOWN_1.
DR PROSITE; PS00061; ETS_DOMAIN_3; 1.
SQ SEQUENCE 428 AA; 47915 MW; 58FEF8ADB985A83 CRC64;

Query Match 4.3%; Score 8; DB 4; Length 428;
Best Local Similarity 100.0%; Pred. No. 6.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 PQGRKPR 171
DB 310 PQGRKPR 317

Mon Mar 31 09:44:34 2003

CC SPASTIC PARALYSIS OF THE ANTERIOR AND POSTERIOR EXTREMITIES AND
 CC DEATH AT DOSE LEVELS OF 0.24 MG/MOUSE. INSECTICIDAL TO THE LARVAL
 CC AND ADULT FORMS OF THE HOUSE FLY.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: Expressed by the venom gland.
 CC -!- SIMILARITY: BELONGS TO THE SPIDER TOXIN TX2 FAMILY.
 DR PIR; B39305; B39305.
 DR PIR; S29215; S29215.
 KW Sodium channel inhibitor; Toxin; Neurotoxin.
 SQ SEQUENCE 49 AA; 5111 MW; 77B46AAB3911716 CRC64;

Query Match 3.2%; Score 6; DB 1; Length 49;
 Best Local Similarity 100.0%; Pred. No. 18;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 138 ASCKCK 143
 |
 |
 |
 |
 |
 Db 44 ASCKCK 49

Search completed: March 28, 2003, 14:22:22
 Job time : 8.89333 secs

CC OF MYOSIN I TO CROSS-LINK ACTIN FILAMENTS.
CC -!- MISCELLANEOUS: THIS ORGANISM EXPRESSES AT LEAST THREE ISOFORMS OF
CC MYOSIN I HEAVY-CHAIN, ENCODED BY GENES MIA, MIB, AND MIC.
CC -!- SIMILARITY: BELONGS TO THE MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC -!- CAUTION: WAS ORIGINALLY THOUGHT TO BE MYOSIN IB.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; J02974; AAA27707.1; -
CC PIR; A33891; MWAXIC.
CC HSP; P08799; 1MND.
CC InterPro; IPR001452; SH3.
CC InterPro; IPR001609; myosin_head.
CC Pfam; PF00018; SH3; 1.
CC Pfam; PF00063; myosin_head; 1.
CC PRINTS; PR00193; MYOSINHEAVY.
CC PRINTS; PR00452; SH3DOMAIN.
CC ProDom; PD000066; SH3; 1.
CC ProDom; PD000353; myosin_head; 1.
CC SMART; SM00242; MYSC; 1.
CC SMART; SM00326; SH3; 1.
CC PROSITE; PS50002; SH3; 1.
CC MYOSIN; ATP-binding; Phosphorylation; Multigene family; SH3 domain.
CC DOMAIN 1 670
CC MYOSIN HEAD-LIKE.
CC DOMAIN 671 922
CC TAIL HOMOLGY REGION 1 (TH.1).
CC DOMAIN 923 975
CC GLY/PRO/ALA-RICH (TH.2).
CC DOMAIN 976 1035
CC SH3.
CC DOMAIN 1036 1168
CC GLY/PRO/ALA-RICH (TH.2).
CC NP_BIND 101 108
CC ATP (POTENTIAL).
CC MOD_RES 311 311
CC PHOSPHORYLATION.
CC SEQUENCE 1168 AA; 127309 MW; D07084B373A37A32 CRC64;

Query Match 3.8%; Score 7; DB 1; Length 1168;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 124 PGGAAPR 130
DB 1103 PGGAAPR 1109
|||||
RESULT 14
BCC3 ACXY
ID BCC3 ACXY STANDARD; PRT; 1325 AA.
AC Q9WJ63;
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Cellulose synthase 1 operon protein C precursor.
BSCL.
OS Acetobacter xylinus.
OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
OC Gluconacetobacter.
OX NCBI_TaxID=28448;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=JCM 7664 / IFO 13693;
RA Umeda Y., Hirano A., Ishibashi M., Akiyama H., Onizuka T., Ikeuchi M.,
RT "Cloning of cellulose synthase genes from Acetobacter xylinum JCM
RT 7664: implication of a novel set of cellulose synthase genes.";
RL DNA Res. 6:109-115(1999).
CC -!- FUNCTION: Required for maximal bacterial cellulose synthesis. It
CC may be involved in the formation of a membrane complex for
CC extrusion of the cellulose product (By similarity).

CC -!- PATHWAY: Bacterial cellulose biosynthesis.
CC -!- SUBCELLULAR LOCATION: Outer membrane (Potential).
CC -!- SIMILARITY: BELONGS TO THE ACSC/BCSC FAMILY.
CC -!- SIMILARITY: CONTAINS 8 TPR REPEATS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AB015802; BAA77587.1; -
CC InterPro; IPR003921; Cellsynth_C.
CC InterPro; IPR001440; TPR.
CC Pfam; PF00515; TPR; 4.
CC PRINTS; PR01441; CELLSYNTHASEC.
CC Cellulose biosynthesis; Repeat; TPR repeat; Membrane; Outer membrane;
KW SIGNAL.
FT CHAIN 1 30
FT REPEAT 31 1325
FT REPEAT 50 33
FT REPEAT 85 117
FT REPEAT 292 325
FT REPEAT 326 359
FT REPEAT 406 439
FT REPEAT 558 591
FT REPEAT 702 735
FT REPEAT 737 769
FT REPEAT 1325 AA; 142541 MW; DA4EECFADDED74D CRC64;
SQ SEQUENCE 1325 AA; 142541 MW; DA4EECFADDED74D CRC64;

Query Match 3.8%; Score 7; DB 1; Length 1325;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 36 NGRPPH 42
DB 161 NGRPPH 167
|||||
RESULT 15
TX25 PHONI
ID TX25 PHONI STANDARD; PRT; 49 AA.
AC P29424;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Neurotoxin Tx2-5.
OS Phoneytria nigriventer (Brazilian armed spider).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;
OC Araneomorphae; Entelegynae; Lycosoidea; Ctenidae; Phoneytria.
OX NCBI_TaxID=6918;
RN [1]
RP SEQUENCE
RC TISSUE=Venom;
RX MEDLINE=93011905; PubMed=1397265;
RA Cordeiro M.N., Diniz C.R., Valentim A.C., von Eickstedt V.R.D.,
RA Gilroy J., Richardson M.;
RT "The purification and amino acid sequences of four Tx2 neurotoxins
RT from the venom of the Brazilian 'armed' spider Phoneytria nigriventer
RT (Keys).";
RL FEBS Lett. 310:153-156(1992).
RN [2]
RP SEQUENCE OF 1-10.
RC TISSUE=Venom;
RX MEDLINE=92196803; PubMed=1801316;
RA Rezende L. Jr., Cordeiro M.N., Oliveira E.B., Diniz C.R.;
RT "Isolation of neurotoxic peptides from the venom of the 'armed'
RT spider Phoneytria nigriventer.";
RL Toxicon 29:1225-1233(1991).
CC -!- FUNCTION: BLOCKS VOLTAGE-GATED SODIUM CHANNELS. CAUSES SCRATCHING,
CC LACRIMATION, HYPERSALIVATION, SWEATING AND AGITATION FOLLOWED BY

FT	CHAIN	35	832	THR-RICH.
FT	DOMAIN	789	813	POTENTIAL.
FT	TRANSMEM	829	849	N-LINKED (GALNAC. . .)
FT	CARBOHYD	36	36	N-LINKED (GLC. . .)
FT	CARBOHYD	339	339	N-LINKED (GLC. . .)
FT	CARBOHYD	338	398	N-LINKED (GLC. . .)
FT	CARBOHYD	438	438	N-LINKED (GLC. . .)
FT	CARBOHYD	513	513	N-LINKED (GLC. . .)
FT	CARBOHYD	643	643	N-LINKED (GLC. . .)

88

RA Olson E.R., Duniak D.S., Jurss L.M., Poorman R.A.;
 RT "Identification and characterization of dppA, an *Escherichia coli*
 RL gene encoding a periplasmic dipeptide transport protein."; J. Bacteriol. 173:234-244(1991).
 RN [2]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 29-50.
 RC STRAIN=K12;
 RX MEDLINE=92065799; PubMed=1956284;
 RA Abouhamad W.N., Manson M., Gibson M.M., Higgins C.F.;
 RT "Peptide transport and chemotaxis in *Escherichia coli* and *Salmonella*
 RL typhimurium: characterization of the dipeptide permease (Dpp) and the
 RN dipeptide-binding protein."; Mol. Microbiol. 5:1035-1047(1991).
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MM500;
 RX MEDLINE=95231288; PubMed=7536291;
 RA Abouhamad W.N., Manson M.D.;
 RT "The dipeptide permease of *Escherichia coli* closely resembles other
 RL bacterial transport systems and shows growth-phase-dependent
 RN expression."; Mol. Microbiol. 14:1077-1092(1994).
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=94316500; PubMed=8041620;
 RA Sofia H.J., Burland V., Daniels D.L., Plunkett G. III, Blattner F.R.;
 RT "Analysis of the *Escherichia coli* genome. V. DNA sequence of the
 RL region from 76.0 to 81.5 minutes."; Nucleic Acids Res. 22:2576-2586(1994).
 RN [5]
 RP SEQUENCE OF 29-40.
 RC STRAIN=K12 / EWG2;
 RX MEDLINE=97443975; PubMed=9298646;
 RA Link A.J., Robison K., Church G.M.;
 RT "Comparing the predicted and observed properties of proteins encoded
 RL in the genome of *Escherichia coli* K-12."; Electrophoresis 18:1259-1313(1997).
 RN [6]
 RP DISULFIDE BONDS.
 RA Duntzen P.;
 RL Submitted (JUN-1995) to the SWISS-PROT data bank.
 RN [7]
 RP X-RAY CRYSTALLOGRAPHY (3.2 ANGSTROMS).
 RX MEDLINE=96155055; PubMed=8563629;
 RA Duntzen P., Mowbray S.L.;
 RT "Crystal structure of the dipeptide binding protein from *Escherichia*
 RL coli involved in active transport and chemotaxis."; Protein Sci. 4:2327-2334(1995).
 RN [8]
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
 RX MEDLINE=96118375; PubMed=8527431;
 RA Nickitenko A.V., Trakhanov S., Quitocho F.A.;
 RT "2-A resolution structure of DppA, a periplasmic dipeptide
 RL transport/chemosensory receptor."; Biochemistry 34:16585-16595(1995).
 CC -!- FUNCTION: DIPEPTIDE-BINDING PROTEIN OF AN OSMOTIC-SHOCKABLE
 CC TRANSPORT SYSTEM. DPPA IS ALSO REQUIRED FOR PEPTIDE CHEMOTAXIS.
 CC -!- SUBCELLULAR LOCATION: Periplasmic.
 CC -!- SIMILARITY: BELONGS TO THE BACTERIAL EXTRACELLULAR SOLUTE-BINDING
 CC PROTEIN FAMILY 5.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M35045; AAA23707.1; -;
 CC EMBL; X58051; CAA41090.1; -;
 CC EMBL; L08399; AAA23702.1; -;

DR EMBL; U00039; AAB18522.1; -;
 DR EMBL; AE000431; AAC76569.1; -;
 DR PIR; S15292; S15292;
 DR PIR; A39194; A39194;
 DR PDB; 1DPE; 17-AUG-96.
 DR PDB; 1DPP; 07-DEC-95.
 DR SWISS-2DPAGE; P23847; COLI.
 DR ECO2DBASE; G059.9; 6TH EDITION.
 DR EcoGene; EG10248; dppA.
 DR InterPro; IPR000914; ssp_bac_5.
 DR Pfam; PF00496; ssp_bac_5_1_5.
 DR PROSITE; PS01040; SBP_BACTERIAL_5; 1.
 KW Peptide transport; Transport; Periplasmic; Signal; Chemotaxis;
 KW 3D-structure; Complete proteome.
 FT SIGNAL 1 28 PERIPLASMIC DIPEPTIDE TRANSPORT PROTEIN.
 FT CHAIN 29 535
 FT DISULFID 34 262
 FT DISULFID 450 463
 SQ SEQUENCE 535 AA; 60293 MW; 3F7C4756E8A8C2F0 CRC64;
 Query Match 3.8%; Score 7; DB 1; Length 535;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 11 TEVIPGL 17
 Db 72 TEVIPGL 78
 RESULT 11
 VGLG SYNV STANDARD; PRT; 632 AA.
 ID VGLG SYNV
 AC P2727; Rel. 23, Created
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DE 01-JUL-1993 (Rel. 26, Last annotation update)
 DE Spike glycoprotein precursor.
 GN G.
 OS Sonchus yellow net virus (SYNV).
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
 OC Rhabdoviridae; Nucleorhabdovirus.
 OX NCBI_TaxID=11307;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC PV-263;
 RX MEDLINE=92024089; PubMed=1926779;
 RA Goldberg K.B., Modrell B., Hillman B.I., Heaton L.A., Choi T.J.,
 RA Jackson A.O.;
 RT "Structure of the glycoprotein gene of sonchus yellow net virus, a
 RL plant rhabdovirus."; Virology 185:32-38(1991).
 CC -!- FUNCTION: THIS PROTEIN FORMS SPIKES ON THE SURFACE OF THE VIRION.
 CC IT IS RESPONSIBLE BOTH FOR THE BINDING OF THE VIRUS TO SUSCEPTIBLE
 CC HOST CELLS AND FOR INDUCING THE UPTAKE OF THE VIRUS BY THE CELL.
 CC THE INTERACTION BETWEEN THE INTERNAL COMPONENTS OF THE VIRION
 CC AND THE PORTION OF THE GLYCOPROTEIN EXPOSED ON THE CYTOPLASMIC
 CC FACE OF THE PLASMA MEMBRANE PROBABLY DIRECTS ENVELOPMENT AND
 CC VIRUS BUDDING.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; L32603; AAA50384.1; -;
 CC EMBL; M73626; AAA47898.1; -;
 CC PIR; A40776; VGVNSY.
 KW Transmembrane; Envelope protein; Glycoprotein; Signal.
 FT SIGNAL 1 17 POTENTIAL.
 FT CHAIN 18 632 SPIKE GLYCOPROTEIN.

```

CC EMBL; X87257; CA860715.1; -
DR EMBL; Z36939; CA85391.1; -
DR HSSP; P28324; 1BC8.
DR TRANSFAC; T05013; -
DR MGD; MGI:101833; Elk1.
DR InterPro; IPR000418; Ets.
DR InterPro; IPR002341; HSF_ETS.
DR Pfam; PF00178; Ets; 1.
DR PRINTS; PR00454; ETSDOMAIN.
DR SMART; SM00413; ETS; 1.
DR PROSITE; PS00345; ETS DOMAIN 1; 1.
DR PROSITE; PS00346; ETS DOMAIN 2; 1.
DR PROSITE; PS00661; ETS DOMAIN 3; 1.
DR Transcription regulation; Activator; Nuclear protein; DNA-binding;
KW Phosphorylation.
KW DNA_BIND 5 ETS-DOMAIN
FT CONFLICT 133 133 P -> T (IN REF. 2).
SQ SEQUENCE 429 AA; 45243 MW; B61B5B977731D54F CRC64;

Query Match 4.3%; Score 8; DB 1; Length 429;
Best Local Similarity 100.0%; Pred. No. 0.82;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 PQGKRP 171
DB 311 PQGKRP 318

RESULT 8
ID VE7 HPV44 STANDARD; PRT; 97 AA.
AC Q80314;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE E7 protein.
GN E7.
OS Human papillomavirus type 44.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10592;
RN [1]
RP SEQUENCE FROM N.A.
RA Delius H.;
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: E7 PROTEIN HAS BOTH TRANSFORMING AND TRANS-ACTIVATING
CC ACTIVITIES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U31788; AAU79458.1; -
DR EMBL; U31788; AAU79458.1; -
DR InterPro; IPR000148; Papi_E7.
DR Pfam; PF00527; E7; 1.
DR Early protein; Transcription regulation; Oncogene;
KW DNA-binding; Trans-acting factor.
FT SITE 57 60 C-XX-C MOTIF-1.
FT SITE 90 93 C-XX-C MOTIF-2.
FT SITE 97 97 C-XX-C MOTIF-3.
SQ SEQUENCE 97 AA; 10641 MW; B4866AE13B050456 CRC64;

Query Match 3.8%; Score 7; DB 1; Length 97;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 131 SRKRLV 137
DB 61 SRKRLV 67

```

```

RESULT 9
ID RL22 LEPIN STANDARD; PRT; 110 AA.
AC Q9XD31;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE 50S ribosomal protein L22.
GN RPLV.
OS Leptospira interrogans.
OC Bacteria; Spirochaetales; Leptospiraceae; Leptospira.
OX NCBI_TaxID=173;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Serovar Lai;
RC MEDLINE=20088835; PubMed=10620683;
RA Zuercher R.L., Hartskeerl R.A., van de Kemp H., Bal A.E.;
RA "Characterization of the Leptospira interrogans S10-spc-alpha
operon.";
RT FEMS Microbiol. Lett. 182:303-308(2000).
CC -1- FUNCTION: THIS PROTEIN BINDS SPECIFICALLY TO 23S rRNA; ITS BINDING
CC IS STIMULATED BY OTHER RIBOSOMAL PROTEINS, E.G., L4, L17, AND L20.
CC IT IS IMPORTANT DURING THE EARLY STAGES OF 50S RECONSTITUTION
CC (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE L22P FAMILY OF RIBOSOMAL PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF115283; AAD40588.1; -
DR HSSP; P48286; IBXE.
DR InterPro; IPR001063; Ribosomal_L22.
DR Pfam; PF00237; Ribosomal_L22; 1.
DR ProDom; PD001032; Ribosomal_L22; 1.
DR TIGRPFAMs; TIGR01044; rplv_bact; 1.
DR PROSITE; PS00464; RIBOSOMAL_L22; 1.
KW Ribosomal protein; rRNA-binding
SQ SEQUENCE 110 AA; 12566 MW; B9181B85E9756CAA CRC64;

Query Match 3.8%; Score 7; DB 1; Length 110;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 132 RKVRLVA 138
DB 15 RKVRLVA 21

RESULT 10
ID DPPA ECOLI STANDARD; PRT; 535 AA.
AC P23847;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Periplasmic dipeptide transport protein precursor (dipeptide-binding
DE protein) (DBP).
GN DPPA OR B3544.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RC MEDLINE=91100289; PubMed=1702779;

```

Db 90 TDGCRSAKPVTELVCGCGPARLLPNAIGR 121

RESULT 6

ELK1_HUMAN

ID ELK1_HUMAN STANDARD; PRT; 428 AA.

AC P19419; O75606; O9UQM4; O95058;

DT 01-NOV-1990 (Rel. 16, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE ETS-domain protein ELK-1.

GN ELK1

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=89203250; PubMed=2539641;

RA Rao V.N., Huebner K., Isobe M., Ar-Rushdi A., Croce C.M.,

RA Reddy E.S.P.;

RT "elk, tissue-specific ets-related genes on chromosomes X and 14 near

RT translocation breakpoints.";

RL Science 244:66-70(1989).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=99013876; PubMed=9795224;

RA Harindranath N., Mills F.C., Mitchell M.P., Meindl A., Max E.E.;

RT "The human elk-1 gene family: the functional gene and two processed

RL pseudogenes embedded in the Igh locus.";

RL Gene 221:215-224(1998).

RN [3]

RP SEQUENCE FROM N.A.

RX Medline=92334979; PubMed=1630903;

RA Janknecht R., Nordheim A.;

RT "Elk-1 protein domains required for direct and SRF-assisted

RT DNA-binding.";

RL Nucleic Acids Res. 20:3317-3324(1992).

CC -!- FUNCTION: STIMULATES TRANSCRIPTION. BINDS TO PURINE-RICH DNA

CC SEQUENCES. CAN FORM A TERNARY COMPLEX WITH THE SERUM RESPONSE

CC FACTOR AND THE ETS AND SRF MOTIFS OF THE FOS SERUM RESPONSE

CC ELEMENT.

CC -!- SUBCELLULAR LOCATION: Nuclear.

CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 (SHOWN HERE) AND 2/ELKV; ARE

CC PRODUCED BY ALTERNATIVE SPLICING.

CC -!- TISSUE SPECIFICITY: LUNG AND TESTIS.

CC -!- PTM: PHOSPHORYLATED BY THE THREE GROUPS OF MAP KINASE (P38, JNK,

CC AND ERK).

CC -!- SIMILARITY: BELONGS TO THE ETS FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

DR EMBL; M25269; AAA52384.1; --

DR EMBL; AF080616; AAC82466.1; --

DR EMBL; AL009172; CAAL15659.1; --

DR EMBL; AF000672; AAD00862.1; --

DR F01; A41354; TVHUEK.

DR HSP: P28324; 1BC8.

DR TRANSFAC; T00250; --

DR Genew; HGNC:3321; ELK1.

DR MIM; 311040; --

DR InterPro; IPR000418; Ets.

DR InterPro; IPR002341; HSF_ETS.

DR Pfam; PF00178; Ets; 1.

DR PRINTS; PR00454; ETSDOMAIN.

DR SMART; SM00413; ETS; 1.

DR PROSITE; PS00345; ETS_DOMAIN_1; 1.

DR PROSITE; PS00346; ETS_DOMAIN_2; 1.

DR PROSITE; PS00611; ETS_DOMAIN_3; 1.

DR Transcription regulation; Activator; Nuclear protein; DNA-binding;

KW Phosphorylation; Alternative splicing.

FT DNA BIND 5 86 ETS-DOMAIN.

FT VARSPLIC 91 95 VAGCS -> SHCAP (IN ISOFORM 2).

FT VARSPLIC 96 428 MISSING (IN ISOFORM 2).

FT CONFLICT 183 193 S -> N (IN REF. 1).

SQ SEQUENCE 428 AA; 44888 MW; 68F71F8ADB9D38CA CRC64;

Query Match 4.3%; Score 8; DB 1; Length 428;

Best Local Similarity 100.0%; Pred. No. 0.82;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 POKGRKPR 171

Db 310 POKGRKPR 317

RESULT 7

ELK1_MOUSE

ID ELK1_MOUSE STANDARD; PRT; 429 AA.

AC P41969;

DT 01-NOV-1995 (Rel. 32, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE ETS-domain protein ELK-1.

GN ELK1

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RX STRAIN=C57BL/6; TISSUE=Embryo;

RX MEDLINE=97017146; PubMed=8663747;

RA Grevin D., Ung S., Denhez F., Dehem M., Quatannens B., Begue A.,

RA Stehelin D., Martin P.;

RT "Structure and organization of the mouse elk1 gene.";

RL Gene 174:185-188(1996).

RN [2]

RP SEQUENCE OF 5-224 FROM N.A.

RC TISSUE=Embryo;

RX MEDLINE=95047310; PubMed=7958835;

RA Glovane A., Pintzas A., Maira S.-M., Sobieszczuk P., Wasylyk B.;

RT "Net, a new ets transcription factor that is activated by Ras.";

RL Genes Dev. 8:1502-1513(1994).

CC -!- FUNCTION: STIMULATES TRANSCRIPTION. BINDS TO PURINE-RICH DNA

CC SEQUENCES. CAN FORM A TERNARY COMPLEX WITH THE SERUM RESPONSE

CC FACTOR AND THE ETS AND SRF MOTIFS OF THE FOS SERUM RESPONSE

CC ELEMENT.

CC -!- SUBCELLULAR LOCATION: Nuclear.

CC -!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN THE BRAIN, AND TO A

CC LESSER EXTENT IN THE HEART, LIVER AND MUSCLE.

CC -!- SIMILARITY: BELONGS TO THE ETS FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

Mon Mar 31 09:44:34 2003

us-09-867-274-4.oligo.rsp

```

FT CARBOHYD 144 144 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT NON_TER 176 176
SQ SEQUENCE 176 AA; 19743 MW; 35F0CA61A425F4DB CRC64;

Query Match 17.3%; Score 32; DB 1; Length 176;
Best Local Similarity 100.0%; Pred. No. 1.4e-27; Indels 0; Gaps 0;
Matches 32; Conservative 0; Mismatches 0;

Qy 65 TDGCRSAKPVTCLVSCGQCPCARLLPNAIGR 96
Db 59 TDGCRSAKPVTCLVSCGQCPCARLLPNAIGR 90

RESULT 4
SOST_CERAE STANDARD; PRT; 213 AA.
AC Q9BQ84;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Sclerostin precursor.
GN SOST.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Cercopithecus.
OX NCBI_TaxID=9534;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21090529; PubMed=11179006;
RA Brunkow M.E., Gardner J.C., Van Ness J., Paepser B.W., Kovacevich B.R.,
RA Prohl S., Skonier J.E., Zhao L., Sabo P.J., Fu Y.H., Alisch R.S.,
RA Gillett L., Colbert T., Tacconi P., Galas D., Hamersma H.,
RA Beighton P., Mulligan J.T.;
RT "Bone dysplasia sclerosteosis results from loss of the SOST gene
product, a novel cysteine knot-containing protein.";
RL Am. J. Hum. Genet. 68:577-589(2001).
CC -!- FUNCTION: Seems to play a role in bone homeostasis (By
CC -!- similarity).
CC -!- SUBCELLULAR LOCATION: Secreted (Potential).
CC -!- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTINE KNOT-LIKE (CTCK) DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF326742; AAK13457.1; -.
CC InterPro: IPR000359; Cys_knot.
CC SMART; SM00041; CT; 1.
CC PROSITE; PS01185; CTCK_1; FALSE_NEG.
CC PROSITE; PS01225; CTCK_2; FALSE_NEG.
CC Signal; Glycoprotein.
CC CHAIN 1 23 POTENTIAL.
CC DOMAIN 24 213 SCLEROSTIN.
CC FT CARBOHYD 82 172 CTCK.
CC FT CARBOHYD 53 53 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 175 175 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC SEQUENCE 213 AA; 23908 MW; 6DA7B5EDE674728A CRC64;

Query Match 17.3%; Score 32; DB 1; Length 213;
Best Local Similarity 100.0%; Pred. No. 1.6e-27; Indels 0; Gaps 0;
Matches 32; Conservative 0; Mismatches 0;

Qy 65 TDGCRSAKPVTCLVSCGQCPCARLLPNAIGR 96
Db 90 TDGCRSAKPVTCLVSCGQCPCARLLPNAIGR 121

RESULT 5

```

```

SOST_HUMAN STANDARD; PRT; 213 AA.
ID Q9BQ84;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Sclerostin precursor.
GN SOST.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=11181578;
RA Balenans W., Ebeling M., Patel N., van Hul E., Olson P., Droszegi M.,
RA Iacza C., Wuyts W., van den Ende J., Willems P., Paes-Alves A.F.,
RA Hill S., Bueno M., Ramos F.J., Tacconi P., Dikkers F.G., Stratakis C.,
RA Lindpaintner K., Vickery B., Foerzler D., Van Hul W.;
RT "Increased bone density in sclerosteosis is due to the deficiency of a
novel secreted protein (SOST).";
RL Hum. Mol. Genet. 10:537-543(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX PubMed=11179006;
RA Brunkow M.E., Gardner J.C., Van Ness J., Paepser B.W., Kovacevich B.R.,
RA Prohl S., Skonier J.E., Zhao L., Sabo P.J., Fu Y.H., Alisch R.S.,
RA Gillett L., Colbert T., Tacconi P., Galas D., Hamersma H.,
RA Beighton P., Mulligan J.T.;
RT "Bone dysplasia sclerosteosis results from loss of the SOST gene
product, a novel cysteine knot-containing protein.";
RL Am. J. Hum. Genet. 68:577-589(2001).
CC -!- FUNCTION: Seems to play a role in bone homeostasis.
CC -!- SUBCELLULAR LOCATION: Secreted (Potential).
CC -!- TISSUE SPECIFICITY: Widely expressed at low levels with highest
CC levels in bone, cartilage, kidney, liver, bone marrow and primary
CC osteoblasts differentiated for 21 days.
CC -!- DISEASE: Defects in SOST are the cause of sclerosteosis, a
CC progressive sclerosing bone dysplasia.
CC -!- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTINE KNOT-LIKE (CTCK) DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF331844; AAK16158.1; -.
CC EMBL; AF326736; AAK13451.1; -.
CC EMBL; AF326739; AAK13454.1; -.
CC Genew; HGNC:13771; SOST.
CC MIM; 605740; -.
CC MIM; 269500; -.
CC InterPro: IPR000359; Cys_knot.
CC SMART; SM00041; CT; 1.
CC PROSITE; PS01185; CTCK_1; FALSE_NEG.
CC PROSITE; PS01225; CTCK_2; FALSE_NEG.
CC Signal; Glycoprotein.
CC CHAIN 1 23 POTENTIAL.
CC DOMAIN 24 213 SCLEROSTIN.
CC FT CARBOHYD 82 172 CTCK.
CC FT CARBOHYD 53 53 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 175 175 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC SEQUENCE 213 AA; 24030 MW; 30BBD55CE73D5BB2 CRC64;

Query Match 17.3%; Score 32; DB 1; Length 213;
Best Local Similarity 100.0%; Pred. No. 1.6e-27; Indels 0; Gaps 0;
Matches 32; Conservative 0; Mismatches 0;

Qy 65 TDGCRSAKPVTCLVSCGQCPCARLLPNAIGR 96

```

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

EMBL; AF326741; AAK13455.1; -
 InterPro; IPR000359; Cys_knot.
 PROSITE; PS01185; CTCK_1; FALSE_NEG.
 Signal; Glycoprotein.
 FT SIGNAL 1 23 POTENTIAL.
 FT CHAIN 24 211 SCLEROSTIN.
 FT DOMAIN 80 170 CTCK.
 FT CARBOHYD 51 51 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 173 173 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 72 72 G -> D (IN REF. 2).
 SQ SEQUENCE 211 AA; 23443 MW; AEB094E358E34961 CRC64;

Query Match 67.6%; Score 125; DB 1; Length 211;
 Best Local Similarity 100.0%; Pred. No. 6e-130;
 Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 VSEYSCRELHYTRLTDPGCRSAKPVTTELVCSCGCGPARLLPNAIGRVKWRPNPDPFRC 109
 DB 73 VSEYSCRELHYTRLTDPGCRSAKPVTTELVCSCGCGPARLLPNAIGRVKWRPNPDPFRC 132
 QY 110 IPDYYAQRVQLLCPGGAAPRSKRVLRVASCCKRLTRFHNOSELKDFGPETARPOKGRK 169
 DB 133 IPDYYAQRVQLLCPGGAAPRSKRVLRVASCCKRLTRFHNOSELKDFGPETARPOKGRK 192
 QY 170 PRGGA 174
 DB 193 PRGGA 197

RESULT 2
 SOST RAT
 ID -SOST RAT STANDARD; PRT; 213 AA.
 AC Q99F67;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Sclerostin precursor.
 GN SOST.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley;
 RX MEDLINE=21090529; PubMed=11179006;
 RA Brunkow M.E., Gardner J.C., Van Ness J., Paepers B.W., Kovacevich B.R.,
 RA Proll S., Skonier J.E., Zhao L., Sabo P.J., Fu Y.H., Alisch R.S.,
 RA Gillett L., Colbert T., Tacconi P., Galas D., Hamersma H.,
 RA Beighton P., Mulligan J.T.;
 RT "Bone dysplasia sclerosteosis results from loss of the SOST gene product, a novel cysteine knot-containing protein.";
 RL Am. J. Hum. Genet. 68:577-589(2001).
 CC -!- FUNCTION: Seems to play a role in bone homeostasis (By similarity).
 CC -!- SUBCELLULAR LOCATION: Secreted (Potential).
 CC -!- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTINE KNOT-LIKE (CTCK) DOMAIN.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

or send an email to license@isb-sib.ch).

EMBL; AF326741; AAK13456.1; -
 InterPro; IPR000359; Cys_knot.
 PROSITE; PS01185; CTCK_1; FALSE_NEG.
 Signal; Glycoprotein.
 FT SIGNAL 1 28 POTENTIAL.
 FT CHAIN 29 213 SCLEROSTIN.
 FT DOMAIN 82 172 CTCK.
 FT CARBOHYD 53 53 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 175 175 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 213 AA; 23974 MW; 6C56C878BCBCD684B CRC64;

Query Match 58.4%; Score 108; DB 1; Length 213;
 Best Local Similarity 100.0%; Pred. No. 3.2e-111;
 Matches 108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 65 TDGFCRSAPVTTELVCSCGCGPARLLPNAIGRVKWRPNPDPFRCIPDRYRAQRVQLLCP 124
 DB 90 TDGFCRSAPVTTELVCSCGCGPARLLPNAIGRVKWRPNPDPFRCIPDRYRAQRVQLLCP 149
 QY 125 GGAAPRSKRVLRVASCCKRLTRFHNOSELKDFGPETARPOKGRKRP 172
 DB 150 GGAAPRSKRVLRVASCCKRLTRFHNOSELKDFGPETARPOKGRKRP 197

RESULT 3

SOST BOVIN
 ID -SOST BOVIN STANDARD; PRT; 176 AA.
 AC Q9BG79;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Sclerostin (Fragment).
 GN SOST.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21090529; PubMed=11179006;
 RA Brunkow M.E., Gardner J.C., Van Ness J., Paepers B.W., Kovacevich B.R.,
 RA Proll S., Skonier J.E., Zhao L., Sabo P.J., Fu Y.H., Alisch R.S.,
 RA Gillett L., Colbert T., Tacconi P., Galas D., Hamersma H.,
 RA Beighton P., Mulligan J.T.;
 RT "Bone dysplasia sclerosteosis results from loss of the SOST gene product, a novel cysteine knot-containing protein.";
 RL Am. J. Hum. Genet. 68:577-589(2001).
 CC -!- FUNCTION: Seems to play a role in bone homeostasis (By similarity).
 CC -!- SUBCELLULAR LOCATION: Secreted (Potential).
 CC -!- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTINE KNOT-LIKE (CTCK) DOMAIN.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

EMBL; AF326738; AAK13453.1; -
 InterPro; IPR000359; Cys_knot.
 SMART; SM00041; CT; 1
 PROSITE; PS01185; CTCK_1; FALSE_NEG.
 PROSITE; PS01225; CTCK_2; FALSE_NEG.
 KW Glycoprotein.
 FT NON TER 1 1
 FT DOMAIN 51 141 CTCK.
 FT CARBOHYD 22 22 N-LINKED (GLCNAC. . .) (POTENTIAL).

GenCore version 5.1.4 p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 28, 2003, 14:18:57 ; Search time 7.89333 Seconds
(without alignments)
972.101 Million cell updates/sec

Title: US-09-867-274-4
Perfect score: 185
Sequence: 1 QGQAFRNDATVDPGLGEY.....KGRKPRGAKANQALENAY 185

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 112892 seqs, 41476328 residues

Word size : 0

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	125	67.6	211	1 SOST_MOUSE	Q99P68 mus musculus
2	108	58.4	213	1 SOST_RAT	Q99P67 rattus norv
3	32	17.3	176	1 SOST_BOVIN	Q99G79 bos taurus
4	32	17.3	213	1 SOST_CERAE	Q99G78 cercopithec
5	32	17.3	213	1 SOST_HUMAN	Q99G74 homo sapien
6	8	4.3	428	1 ELK1_HUMAN	P19419 mus musculu
7	8	4.3	429	1 ELK1_MOUSE	Q80914 human papil
8	7	3.8	97	1 VE7_HPV44	Q80914 leptospira
9	7	3.8	110	1 RL22_LEPIN	P23847 escherichia
10	7	3.8	535	1 DPBA_ECOLI	P27277 sonchus yel
11	7	3.8	832	1 VGLG_SVNV	P08198 halobacteri
12	7	3.8	852	1 CSG_HAHLA	P10569 acanthamoeb
13	7	3.8	1168	1 MYSC_ACACA	Q99X63 acetobactria
14	7	3.8	1325	1 BCC3_ACEXY	P29424 phonetria
15	6	3.2	49	1 TX25_PHONI	P07887 corynebacte
16	6	3.2	107	1 TH12_CORNE	P02423 escherichia
17	6	3.2	110	1 RL22_ECOLI	P29222 acholeplasma
18	6	3.2	111	1 RL22_ACHLA	O31160 spiroplasma
19	6	3.2	112	1 RL22_SPICI	P48286 thermus the
20	6	3.2	113	1 RL22_TRETH	O95411 homo sapien
21	6	3.2	115	1 T1AF_HUMAN	Q92184 mus musculu
22	6	3.2	115	1 T1AF_MOUSE	Q92184 mus musculu
23	6	3.2	118	1 REV_HVILW	Q79624 human immun
24	6	3.2	121	1 RL7_LACLA	O95942 lactococcus
25	6	3.2	131	1 RL22_PHYS1	O66094 phytoplasma
26	6	3.2	166	1 RS10_ICTPU	Q90YV4 ictalurus p
27	6	3.2	167	1 B3AR_MERUN	O70432 meriones un
28	6	3.2	173	1 FRIS_LYNST	P42577 lymnaea sta
29	6	3.2	182	1 YCYO_YEAST	P25654 saccharomyc
30	6	3.2	193	1 HS72_CANAL	P46587 candida alb
31	6	3.2	195	1 YL47_ARCFU	O28135 archaeoglob
32	6	3.2	196	1 RETB_CHICK	P41263 gallus gall
33	6	3.2	198	1 RECR_STRPY	Q99Z33 streptococc

34 6 3.2 210 1 TRPF_KLUULA
35 6 3.2 213 1 PYRE_HABIN
36 6 3.2 214 1 NADD_PSEAE
37 6 3.2 216 1 FGFH_HUMAN
38 6 3.2 216 1 FGFH_MOUSE
39 6 3.2 221 1 KGUW_RALSO
40 6 3.2 227 1 PRL_HUMAN
41 6 3.2 227 1 PRL_MACMU
42 6 3.2 230 1 COX2_BRAFL
43 6 3.2 239 1 COX2_BRALA
44 6 3.2 249 1 APX1_PEA
45 6 3.2 261 1 DHSB_RICPR

ALIGNMENTS

RESULT 1
SOST_MOUSE STANDARD; PRT; 211 AA.
ID Q99P68; Q99D17;
AC 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Sclerostin precursor.
GN SOST.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=129/SVJ;
RX MEDLINE=21090529; PubMed=11179006;
RA Brunkow M.E., Gardner J.C., Van Ness J., Paepers B.W., Kovacevich B.R.,
RA Proll S., Skonier J.E., Zhao L., Sabo P.J., Fu Y.H., Alisch R.S.,
RA Gillett L., Colbert T., Tacconi P., Galas D., Hamersma H.,
RA Beighton P., Mulligan J.T.;
RT "bone dysplasia sclerosteosis results from loss of the SOST gene
product, a novel cysteine knot-containing protein.";
RL Am. J. Hum. Genet. 68:577-589 (2001).
[2]
SEQUENCE FROM N.A.
RP STRAIN=C57BL/6J; TISSUE=Head;
RC MEDLINE=21085660; PubMed=11217851;
RX Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Stauber P., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hall D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombauer P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schenbach C., Seva T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohetsuki S.,
RA Hasehizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690 (2001).
CC -!- FUNCTION: Seems to play a role in bone homeostasis (By
similarity).
CC -!- SUBCELLULAR LOCATION: Secreted (potential).
CC -!- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTEINE KNOT-LIKE (CTCK) DOMAIN.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its

A:Gene: UU417
A:Genetic code: SGC3

Query Match 3.8%; Score 7; DB 2; Length 321;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 133 KVLVLAS 139
|||||
DB 276 KVLVLAS 282

RESULT 12

S72647

hypothetical protein lb - anthracnose fungus (Colletotrichum gloeosporioides) retrotransposon
C:Species: Glomerella cingulata, Colletotrichum gloeosporioides
C:Date: 24-Oct-1998 #sequence_revision 24-Oct-1998 #text_change 17-Mar-1999
C:Accession: S72647

R:He, C.; Nourse, J.P.; Kelemu, S.; Irwin, J.A.G.; Manners, J.M.
Mol. Gen. Genet. 252, 320-331, 1996

A:Title: Cgtr1: a non-LTR retrotransposon with restricted distribution in the fungal phyto

A:Reference number: 372619; MUID:96439839; PMID:8842152

A:Accession: S72647

A:Status: not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-328 <HEH>

A:Experimental source: biotype B, isolate UQ62

C:Genetics:

A:Mobile element: retrotransposon Cgtr1

Query Match 3.8%; Score 7; DB 2; Length 328;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 85 GPARLLP 91
|||||
DB 123 GPARLLP 129

RESULT 13

C83577

hypothetical protein PA0549 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: C83577

R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
; Loty, S.; Olson, M.V.
Nature 406, 959-964, 2000

A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho

A:Reference number: A82950; MUID:20437337; PMID:10984043

A:Accession: C83577

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-354 <STO>

A:Cross-references: GB:AE004491; GB:AE004091; NID:g9946412; PIDN:AA03938.1; GSPDB:GN001

A:Experimental source: strain PA01

C:Genetics:

A:Gene: PA0549

Query Match 3.8%; Score 7; DB 2; Length 354;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 95 GRVKWR 101
|||||
DB 206 GRVKWR 212

RESULT 14

S24802

polyferredoxin 6x2[4Fe-4S] vhuB [similarity] - Methanococcus voltae
C:Species: Methanococcus voltae

C:Date: 10-Sep-1999 #sequence_revision 18-Aug-2000 #text_change 19-Jan-2001
C:Accession: S24802

R:Halboth, S.; Klein, A.

submitted to the EMBL Data Library, August 1991

A:Description: Methanococcus voltae harbors two gene groups each of homologous (Nife) - ai

A:Reference number: S16721

A:Accession: S24802

A:Molecule type: DNA

A:Residues: 1-398 <HAL>

A:Cross-references: EMBL:X61204; NID:g1747406; PIDN:CAA43512.1; PID:g1747410

A:Experimental source: strain PS(DSM1537)

R:Halboth, S.; Klein, A.

Mol. Gen. Genet. 233, 217-224, 1992

A:Title: Methanococcus voltae harbors four gene clusters potentially encoding two [Nife]

A:Reference number: A59304; MUID:92293118; PMID:1603063

A:Contents: annotation

C:Genetics:

A:Gene: vhuB

C:Superfamily: polyferredoxin 6x2[4Fe-4S]; ferredoxin 2[4Fe-4S] homology

F:4-52/Domain: ferredoxin 2[4Fe-4S] homology <FER1>

F:56-109/Domain: ferredoxin 2[4Fe-4S] homology <FER2>

F:125-179/Domain: ferredoxin 2[4Fe-4S] homology <FER3>

F:192-247/Domain: ferredoxin 2[4Fe-4S] homology <FER4>

F:261-329/Domain: ferredoxin 2[4Fe-4S] homology <FER5>

F:341-395/Domain: ferredoxin 2[4Fe-4S] homology <FER6>

Query Match 3.8%; Score 7; DB 1; Length 398;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 71 SAKPVTE 77
|||||
DB 252 SAKPVTE 258

RESULT 15

C89753

protein Fl1C7.3 [imported] - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Nov-2001

C:Accession: C89753

R:anonymous, The C. elegans Sequencing Consortium.

Science 282, 2012-2018, 1998

A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biology

A:Reference number: A75000; MUID:99069613; PMID:9851916

A:Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C.eleg

A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and

A:Accession: C89753

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-425 <STO>

A:Cross-references: GB:chr_X; PIDN:AAC69015.1; PID:g2914120; GSPDB:GN00028; CESP:Fl1C7.3

C:Genetics:

A:Gene: Fl1C7.3

A:Map position: X

Query Match 3.8%; Score 7; DB 2; Length 425;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 152 SELKDFG 158
|||||
DB 332 SELKDFG 338

Search completed: March 28, 2003, 14:23:59

Job time : 16.3067 secs

QY 172 PGAKANQ 178

Db 6 PGAKANQ 12

RESULT 7

G81057 hypothetical protein NMB1656 [imported] - Neisseria meningitidis (strain MC58 serogroup

C;Species: Neisseria meningitidis

C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001

C;Accession: G81057

R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.

Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;

Xi, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Massignani, V.; Pizzi, M.

Science 287, 1809-1815, 2000

A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; V

A;Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.

A;Reference number: A81000; MUID:20175755; PMID:10710307

A;Accession: G81057

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-197 <TET>

A;Cross-references: GB:AE002516; GB:AE002098; NID:g7226905; PIDN:AAF42005.1; PID:g722690

A;Experimental source: serogroup B, strain MC58

C;Genetics:

A;Gene: NMB1656

Query Match 3.8%; Score 7; DB 2; Length 197;

Best Local Similarity 100.0%; Pred. No. 18;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 175 KANQAEI 181

Db 122 KANQAEI 128

RESULT 8

A81819 hypothetical protein NMA1913 [imported] - Neisseria meningitidis (strain Z2491 serogroup

C;Species: Neisseria meningitidis

C;Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001

C;Accession: A81819

R;Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel

; Holroyd, S.; Jørgensen, K.; Leather, S.; Mouton, K.; Quail, M.A.; Rajandream,

Nature 404, 502-506, 2000

A;Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.

A;Reference number: A81775; MUID:20222556; PMID:10761919

A;Accession: A81819

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-197 <PAR>

A;Cross-references: GB:AL162757; GB:AL157959; NID:g7380371; PIDN:CAB85134.1; PID:g738054

A;Experimental source: serogroup A, strain Z2491

C;Genetics:

A;Gene: NMA1913

Query Match 3.8%; Score 7; DB 2; Length 197;

Best Local Similarity 100.0%; Pred. No. 18;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 175 KANQAEI 181

Db 122 KANQAEI 128

RESULT 9

AI2930 hypothetical protein Atu3047 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)

C;Species: Agrobacterium tumefaciens

C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 01-Feb-2002

C;Accession: AI2930

R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I

erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell

; Karp, P.; Romero, P.; Zhang, S.

Science 294, 2317-2323, 2001

A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, I

ster, E.W.

A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A;Reference number: AB2577; PMID:11743193

A;Accession: AI2930

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-296 <KUR>

A;Cross-references: GB:AE008689; PIDN:AAI43863.1; PID:g17741408; GSPDB:GN00187

A;Experimental source: strain C58 (Dupont)

C;Genetics:

A;Gene: Atu3047

A;Map position: linear chromosome

C;Superfamily: oligopeptide permease protein oppB

Query Match 3.8%; Score 7; DB 2; Length 296;

Best Local Similarity 100.0%; Pred. No. 25;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 89 LLPNAIG 95

Db 206 LLPNAIG 212

RESULT 10

E98351 oligopeptide transport system permease protein appC AGR_L_3516 [imported] - Agrobacterium

C;Species: Agrobacterium tumefaciens

C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 11-Jan-2002

C;Accession: E98351

R;Goodner, B.; Hinkley, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,

A.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.;

Science 294, 2323-2328, 2001

A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum

A;Reference number: A97359; PMID:11743194

A;Accession: E98351

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-311 <KUR>

A;Cross-references: GB:AE007870; PIDN:AAK90335.1; PID:g15160372; GSPDB:GN00170

C;Genetics:

A;Gene: AGR_L_3516

A;Map position: linear chromosome

C;Superfamily: oligopeptide permease protein oppB

Query Match 3.8%; Score 7; DB 2; Length 311;

Best Local Similarity 100.0%; Pred. No. 27;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 89 LLPNAIG 95

Db 221 LLPNAIG 227

RESULT 11

B82892 conserved hypothetical U0417 [imported] - Ureaplasma urealyticum

C;Species: Ureaplasma urealyticum

C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000

C;Accession: B82892

R;Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.

submitted to GenBank, February 2000

A;Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a mini

A;Reference number: A82870

A;Accession: B82892

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-321 <GLA>

A;Cross-references: GB:AE002139; GB:AF222894; NID:56899405; PIDN:AAF30828.1; GSPDB:GN001

A;Experimental source: serovar 3; biovar 1

C;Genetics:

A;Cross-references: EMBL:X87257; NID:g836634; PIDN:CAA60715.1; PID:g836635
A;Experimental source: embryo
R;Giovane, A.; Pintzas, A.; Maira, S.M.; Sobieszczuk, P.; Wasyluk, B.
Genes Dev. 8, 1502-1513, 1994
A;Title: Net, a new ets transcription factor that is activated by Ras.
A;Reference number: A53837; MUID:95047310; PMID:7958835
A;Accession: I48339
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-429 <RES>
A;Cross-references: EMBL:X87257; NID:g836634; PIDN:CAA60715.1; PID:g836635
A;Accession: I48340
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 5-132, "T", 134-224 <RE2>
A;Cross-references: EMBL:Z36939; NID:g535922; PIDN:CAA85391.1; PID:g535923
C;Comment: This protein belongs to the subfamily of ternary complex factor (TCF) which b
C;Genetics:
A;Gene: elk1
A;Introns: 70/3; 219/3; 363/3; 397/3
C;Superfamily: elk-1 transforming protein; ets DNA-binding domain homology
F;7-86/Domain: ets DNA-binding domain homology <ETS>

Query Match 4.3%; Score 8; DB 2; Length 429;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 PQKGRKPR 171
|||||||
Db 311 PQKGRKPR 318

RESULT 3
AG0002
conserved hypothetical protein YP00013a [imported] - Yersinia pestis (strain CO92)
C;Species: Yersinia pestis
C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 02-Nov-2001
C;Accession: AG0002
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
Il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall,
Nature 413, 523-527, 2001
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A;Reference number: AB0001; MUID:21470413; PMID:11586360
A;Accession: AG0002
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-89 <KUR>
A;Cross-references: GB:AL590842; PIDN:CAC88881.1; PID:gi5978129; GSPDB:GN00175
C;Genetics:
A;Gene: YP00013a

Query Match 3.8%; Score 7; DB 2; Length 89;
Best Local Similarity 100.0%; Pred. No. 9.4;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 TEVIPGL 17
|||||||
Db 62 TEVIPGL 68

RESULT 4
D83305
hypothetical protein PA2722 [imported] - Pseudomonas aeruginosa (strain PA01)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C;Accession: D83305
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A;Reference number: AB2950; MUID:20437337; PMID:10984043

A;Accession: D83305
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-130 <STO>
A;Cross-references: GB:AE004700; GB:AE004091; NID:g9948792; PIDN:AAG06110.1; GSPDB:GN001;
A;Experimental source: strain PA01
C;Genetics:
A;Gene: PA2722

Query Match 3.8%; Score 7; DB 2; Length 130;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 133 KVLVLAS 139
|||||||
Db 12 KVLVLAS 18

RESULT 5
CB3091
hypothetical protein PA4441 [imported] - Pseudomonas aeruginosa (strain PA01)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C;Accession: CB3091
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A;Reference number: AB2950; MUID:20437337; PMID:10984043
A;Accession: CB3091
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-148 <STO>
A;Cross-references: GB:AE004858; GB:AE004091; NID:g9950668; PIDN:AAG07829.1; GSPDB:GN001;
A;Experimental source: strain PA01
C;Genetics:
A;Gene: PA4441

Query Match 3.8%; Score 7; DB 2; Length 148;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 87 ARLLPNA 93
|||||||
Db 26 ARLLPNA 32

RESULT 6
AC1832
hypothetical protein alr0203 [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp.
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002
C;Accession: AC1832
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yanada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anat
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AC1832
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-162 <KUR>
A;Cross-references: GB:BA000019; PIDN:BA077727.1; PID:gl7135181; GSPDB:GN00179
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: alr0203

Query Match 3.8%; Score 7; DB 2; Length 162;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

GenCore version 5.1.4_p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 28, 2003, 14:20:22 ; Search time 14.3067 Seconds
(without alignments)
1243.118 Million cell updates/sec

Title: US-09-867-274-4
Perfect score: 185
Sequence: 1 QGQAFRNDATEVIPGLGEY.....KGRKPRPCAKAQAELNAY 195

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283224 seqs, 96134422 residues

Word size : 0
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : PIR 73.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8	4.3	428	1 TVHUEK	transforming prote
2	8	4.3	429	2 JC4965	elki protein - mou
3	7	3.8	89	2 AG0002	conserved hypothet
4	7	3.8	130	2 D83305	hypothetical prote
5	7	3.8	148	2 C83091	hypothetical prote
6	7	3.8	162	2 AC1832	hypothetical prote
7	7	3.8	197	2 G81057	hypothetical prote
8	7	3.8	197	2 A81819	hypothetical prote
9	7	3.8	296	2 AI2930	hypothetical prote
10	7	3.8	311	2 E98351	conserved hypothet
11	7	3.8	321	2 B82892	hypothetical prote
12	7	3.8	328	2 S72647	hypothetical prote
13	7	3.8	354	2 C93577	polyferredoxin 6x2
14	7	3.8	398	1 S24802	protein Fl1C7.3 [i
15	7	3.8	425	2 C89753	probable binding p
16	7	3.8	531	2 B83082	periplasmic dipept
17	7	3.8	535	1 A39194	dipeptide transpor
18	7	3.8	535	2 H91181	dipeptide transpor
19	7	3.8	535	2 D86028	periplasmic dipept
20	7	3.8	542	2 T23869	hypothetical prote
21	7	3.8	562	2 T05758	hypothetical prote
22	7	3.8	565	2 B70652	probable cyto prot
23	7	3.8	632	1 VGVNSY	surface glycoprote
24	7	3.8	632	2 AG1547	transcription anti
25	7	3.8	681	2 E82812	outer membrane hem
26	7	3.8	730	2 B87265	TonB-dependent rec
27	7	3.8	836	2 B84417	cell surface glyco
28	7	3.8	852	2 A28459	cell surface glyco
29	7	3.8			

30 7 3.8 1008 2 B95979 probable enzyme wi
31 7 3.8 1168 1 MWAXIC myosin heavy chain
32 7 3.8 1308 2 T15280 hypothetical prote
33 7 3.8 1488 2 AG2136 polyketide synthas
34 7 3.8 1940 2 AS9287 myosin heavy chain
35 6 3.2 22 2 JF0066 ribosomal protein
36 6 3.2 49 2 S29215 neurotoxin Tx2 - g
37 6 3.2 89 2 D75271 hypothetical prote
38 6 3.2 102 2 G85584 unknown protein en
39 6 3.2 102 2 E90734 hypothetical prote
40 6 3.2 103 2 G84293 hypothetical prote
41 6 3.2 108 1 S02802 thioredoxin C-2 -
42 6 3.2 108 2 F82480 hypothetical prote
43 6 3.2 108 2 D90051 hypothetical prote
44 6 3.2 110 1 R5EC22 ribosomal protein
45 6 3.2 110 2 D91151 50S ribosomal subu

ALIGNMENTS

RESULT 1
TVHUEK
transforming protein elk-1 - human
C;Species: Homo sapiens (man)
C;Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 18-Jun-1999
C;Accession: A41354; S54721
R;Rao, V.N.; Huebner, K.; Isobe, M.; ar-Rushdi, A.; Croce, C.M.; Reddy, E.S.P.
Science 244, 66-70, 1989
A;Title: elk, tissue-specific ets-related genes on chromosomes X and 14 near translocation
A;Reference number: A41354; MUID:89203250; PMID:2539641
A;Accession: A41354
A;Molecule type: mRNA
A;Residues: 1-428 <RAO>
A;Cross-references: GB:M25269; NID:G538208; PIDN:AAA52384.1; PID:G538209
R;Gille, H.; Kortenmann, M.; Thomae, O.; Moomaw, C.; Slaughter, C.; Cobb, M.H.; Shaw, P.F.
EMBO J. 14, 951-962, 1995
A;Title: ERK phosphorylation potentiates Elk-1-mediated ternary complex formation and tra
A;Reference number: S54721; MUID:95196758; PMID:7889942
A;Accession: S54721
A;Status: preliminary
A;Molecule type: protein
A;Residues: 318-328, 'XX', 331, 336-364; 380-388, 'X', 390-392, 'X', 394-400, 'XX', 403-405, 'X', 407
C;Genetics:
A;Gene: GDB:ELK1
A;Cross-references: GDB:119867; OMIM:311040
A;Map position: Xp11.2-Xp11.2
C;Superfamily: elk-1 transforming protein; ets DNA-binding domain homology
C;Keywords: DNA binding; oncogene; transforming protein
F;7-86/Domain: ets DNA-binding domain homology <ETS>

Query Match 4.3%; Score 8; DB 1; Length 428;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 PQGRKPR 171
| | | | |
DB 310 PQGRKPR 317

RESULT 2
JC4965
elki protein - mouse
C;Species: Mus musculus (house mouse)
C;Date: 31-Dec-1996 #sequence_revision 31-Dec-1996 #text_change 16-Jul-1999
C;Accession: JC4965; I48339; I48340; S54908
R;Grevin, D.; Ung, S.; Denhez, F.; Dehem, M.; Quatmanns, B.; Begue, A.; Stehelin, D.; Me
Gene 174, 185-188, 1996
A;Title: Structure and organization of the mouse elki gene.
A;Reference number: JC4965; MUID:97017146; PMID:8863747
A;Accession: JC4965
A;Molecule type: mRNA
A;Residues: 1-429 <GRE>

Mon Mar 31 09:44:32 2003

```

PR 28-FEB-2001; 2001WO-US06520.
PR 09-MAR-2001; 2001US-274399P.
PR 03-APR-2001; 2001US-280982P.
PR 04-APR-2001; 2001US-282129P.
PR 04-APR-2001; 2001US-282199P.
PR 09-MAY-2001; 2001US-290589P.
PR 25-MAY-2001; 2001WO-US17092.
PR 01-JUN-2001; 2001WO-US17800.
PR 20-JUN-2001; 2001WO-US19692.
PR 29-JUN-2001; 2001WO-US21066.
PR 09-JUL-2001; 2001WO-US21735.
XX
XX (GETH ) GENENTECH INC.
XX
XX Baker KP, Eaton DL, Filvaroff E, Goddard A, Grimaldi JC;
PI Gurney AL, Smith V, Stephan J, Watanabe CK, Wood WI, Zhang Z;
PI Fong S;
XX
XX WPI; 2002-362426/39.
DR N-PSDB; ABK69992.
DR
XX New PRO polypeptides and polynucleotides encoding the polypeptides,
PT useful in gene therapy, chromosome identification, tissue typing, or
PT for genetic analysis of individuals with genetic disorders
XX
XX Claim 11; Figure 64; 218pp; English.
XX
XX This invention relates to the cDNA and protein sequences of novel
CC secreted and transmembrane polypeptides PRO polypeptides. The
CC invention also comprises a method for producing the proteins of the
CC invention by recombinant means and antibodies specific for the protein
CC of the invention. The antibody may be used for detecting the PRO
CC proteins of the invention and may be used to modify their activity.
CC Polynucleotides may be used as hybridisation probes for a cDNA library
CC to isolate the full-length PRO cDNA or to isolate other cDNAs, to
CC construct hybridisation probes for mapping the gene which encodes that
CC PRO and for genetic analysis of individuals with genetic disorders, in
CC assays to identify other proteins or molecules involved in binding
CC reaction, to generate transgenic animals or knock-out animals which in
CC turn are useful in the development and screening of therapeutically
CC useful reagents, for chromosome identification, and tissue typing. The
CC PRO polypeptides are useful in gene therapy, and as molecular weight
CC markers for protein electrophoresis purposes. The sequences may
CC also be used to detect overexpression on PRO polypeptides in cancerous
CC tumours and for screening for differentially expressed genes using
CC microarray technology. The present sequence represents a human PRO
CC protein of the invention.
XX
XX Sequence 213 AA;
SQ
Query Match 17.3%; Score 32; DB 23; Length 213;
Best Local Similarity 100.0%; Pred.No. 7.5e-25;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 65 TDGFCRSAPVTELVCSCGCGPARLLPNAIGR 96
DB 90 TDGFCRSAPVTELVCSCGCGPARLLPNAIGR 121

```

Search completed: March 28, 2003, 14:21:58
 Job time : 33.56 secs

DR WPI; 2000-412321/35.
 XX N-PSDB; AAA29062.
 PT Nucleic acids (1) encoding a transforming growth factor beta binding
 PT protein, useful for identifying agents for treating osteopenia,
 PT osteoporosis and fractures
 XX
 PS Disclosure; Page 121; 162pp; English.
 XX
 CC This shows a variant human transforming growth factor-beta
 CC (TGF-beta) binding protein designated BEER P38R. The encoded protein
 CC comprises a substitution of arginine for the wild-type proline at
 CC residue 38. The cDNA and protein may be used for prevention, treatment
 CC and diagnosis of diseases associated with inappropriate BEER expression.
 CC For example, they may be used to treat disorders associated with
 CC decreased TGF-beta BP expression. The cDNA or vectors may be administered
 CC to treat diseases by rectifying mutations or deletions in a patient's
 CC genome that affect the activity of BEER by expressing inactive proteins
 CC or to supplement the patients own production of BEER polypeptides. The
 CC nucleic acids may be used for recombinant production of BEER. Gene
 CC therapy, antisense therapy, as probes for diagnostic assays and for
 CC functional studies. BEER may be used to raise antibodies and for
 CC identification of BEER modulators. BEER antagonists may be used to
 CC increase bone mineral content for the treatment of disorders such as
 CC osteopenia, osteoporosis, fractures and other disorders associated with
 CC low mineral content.
 XX
 SQ Sequence 213 AA;
 Query Match 17.3%; Score 32; DB 21; Length 213;
 Best Local Similarity 100.0%; Pred. No. 7.5e-25;
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 65 TDGFCRSAPVTVELVCSGCGPARLLPNAIGR 96
 DB 90 TDGFCRSAPVTVELVCSGCGPARLLPNAIGR 121
 RESULT 14
 AAAY97589
 ID AAY97589 standard; Protein; 213 AA.
 AC AAY97589;
 XX
 DT 05-APR-2001 (first entry)
 DE Human secreted protein PRO7476.
 XX
 KW Secreted protein; human; PRO protein; neoplastic cell growth; tumour;
 KW proliferation; leukaemia; lymphoid malignancy; inflammatory disorder;
 KW angiogenic disorder; immunologic disorder; PRO7476.
 XX
 OS Homo sapiens.
 XX
 PN WO200075317-A2.
 PD 14-DEC-2000.
 XX
 PF 15-MAY-2000; 2000WO-US13358.
 XX
 PR 09-JUN-1999; 99US-0138385.
 PR 20-JUL-1999; 99US-0144790.
 PR 03-AUG-1999; 99US-0146843.
 PR 10-AUG-1999; 99US-0148188.
 PR 17-AUG-1999; 99US-0149320.
 PR 17-AUG-1999; 99US-0149327.
 PR 17-AUG-1999; 99US-0149396.
 PR 20-AUG-1999; 99US-0150114.
 PR 31-AUG-1999; 99US-0151700.
 PR 31-AUG-1999; 99US-0151734.
 XX
 PA (GETH) GENENTECH INC.
 XX

PI Botstein DA, Goddard A, Gurney AL, Smith V, Watanabe CK, Wood WI;
 XX WPI; 2001-071075/08.
 DR N-PSDB; AAA91023.
 XX
 PT Antibodies against PRO polypeptides, useful for diagnosing and treating
 PT tumours are associated with gene amplification, neoplastic cell growth
 PT and proliferation in mammals -
 XX
 PS Claim 61; Fig 20; 143pp; English.
 XX
 CC This sequence is a human PRO protein of the invention. The PRO
 CC proteins are secreted proteins. Antagonists or antibodies of PRO
 CC polypeptides are useful for diagnosing and treating tumours are
 CC associated with gene amplification, neoplastic cell growth and
 CC proliferation in mammals, and those conditions characterised by
 CC overexpression and/or activation of the amplified genes. Such conditions
 CC include benign or malignant tumours (e.g. renal, liver, kidney, bladder,
 CC breast, gastric, ovarian, colorectal, prostate, pancreatic, lung, vulval,
 CC thyroid, hepatic carcinomas, sarcomas, glioblastomas and various head and
 CC neck tumours); leukemias and lymphoid malignancies; neuronal, glial
 CC astrocytal, hypothalamic, and other glandular, macrophageal, epithelial,
 CC stromal and blastoclastic disorders; and inflammatory, angiogenic and
 CC immunologic disorders. These may further be used to qualitatively or
 CC quantitatively detect the expression of proteins encoded by the
 CC amplified genes, and in tumour diagnostics or prognostics. The PRO
 CC polypeptide or its antagonist may be used for the preparation of a
 CC medicament in the treatment of a condition, which is responsive to the
 CC PRO polypeptide, its antagonist or anti-PRO antibody.
 XX
 SQ Sequence 213 AA;
 Query Match 17.3%; Score 32; DB 22; Length 213;
 Best Local Similarity 100.0%; Pred. No. 7.5e-25;
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 65 TDGFCRSAPVTVELVCSGCGPARLLPNAIGR 96
 DB 90 TDGFCRSAPVTVELVCSGCGPARLLPNAIGR 121
 RESULT 15
 ABG34061
 ID ABG34061 standard; Protein; 213 AA.
 AC ABG34061;
 XX
 DT 15-JUL-2002 (first entry)
 DE Human Pro peptide #32.
 XX
 KW Human; PRO; secreted protein; transmembrane protein;
 KW genetic disorder; tumour; cancer.
 XX
 OS Homo sapiens.
 XX
 PN WO200224888-A2.
 PD 28-MAR-2002.
 XX
 PF 29-AUG-2001; 2001WO-US27099.
 XX
 PR 01-SEP-2000; 2000US-239896P.
 PR 05-SEP-2000; 2000US-230621P.
 PR 22-SEP-2000; 2000US-235147P.
 PR 10-NOV-2000; 2000WO-US30873.
 PR 12-JAN-2001; 2001US-261878P.
 PR 16-JAN-2001; 2001US-261910P.
 PR 16-JAN-2001; 2001US-261939P.
 PR 16-JAN-2001; 2001US-262150P.
 PR 25-JAN-2001; 2001US-264395P.
 PR 02-FEB-2001; 2001US-266421P.
 PR 09-FEB-2001; 2001US-267623P.

Mon Mar 31 09:44:32 2003

XX 24-NOV-1999; 99WO-US27990.
 XX 27-NOV-1998; 98US-0110283.
 XX (DARW-) DARWIN DISCOVERY LTD.
 XX Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepers BW;
 XX Van Ness J, Winkler DG;
 XX WPI; 2000-412321/35.
 XX N-PSDB; AAA29056.
 XX Nucleic acids (I) encoding a transforming growth factor beta binding
 XX protein, useful for identifying agents for treating osteopenia,
 XX osteoporosis and fractures
 XX Claim 3; Page 119-120; 162pp; English.
 XX This shows a variant human transforming growth factor-beta (TGF-beta)
 XX binding protein designated BEER V101, which comprises a substitution of
 XX isoleucine for the wild-type valine at residue 10. The cDNA and protein
 XX may be used for prevention, treatment and diagnosis of diseases
 XX associated with inappropriate BEER expression. For example, they may be
 XX used to treat disorders associated with decreased TGF-beta BP expression.
 XX The cDNA or vectors may be administered to treat diseases by rectifying
 XX mutations or deletions in a patient's genome that affect the activity of
 XX BEER by expressing inactive proteins or to supplement the patients own
 XX production of BEER polypeptides. The nucleic acids may be used for
 XX recombinant production of BEER, gene therapy, antisense therapy, as
 XX probes for diagnostic assays and for functional studies. BEER may be used
 XX to raise antibodies and for identification of BEER modulators. BEER
 XX antagonists may be used to increase bone mineral content for the
 XX treatment of disorders such as osteopenia, osteoporosis, fractures and
 XX other disorders associated with low mineral content.
 XX Sequence 213 AA;
 XX Query Match 17.3%; Score 32; DB 21; Length 213;
 XX Best Local Similarity 100.0%; Pred. No. 7.5e-25;
 XX Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX QY 65 TDGPCRSKAPVTELVCSCGCGPARLLPNAIGR 96
 XX |||||
 XX DB 90 TDGPCRSKAPVTELVCSCGCGPARLLPNAIGR 121
 XX
 XX RESULT 12
 XX AAY96431
 XX ID AAY96431 standard; Protein; 213 AA.
 XX AC AAY96431;
 XX DT 12-SEP-2000 (first entry)
 XX DE Vervet TGF-beta binding protein (BEER).
 XX osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
 XX BEER; gene therapy; antisense therapy; fracture; bone mineralization.
 XX Cercopithecus pygerythrus.
 XX WO200032773-A1.
 XX 08-JUN-2000.
 XX 24-NOV-1999; 99WO-US27990.
 XX 27-NOV-1998; 98US-0110283.
 XX (DARW-) DARWIN DISCOVERY LTD.
 XX Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepers BW;
 XX Van Ness J, Winkler DG;

PI Van Ness J, Winkler DG;
 XX WPI; 2000-412321/35.
 DR N-PSDB; AAA29057.
 XX Nucleic acids (I) encoding a transforming growth factor beta binding
 PT protein, useful for identifying agents for treating osteopenia,
 PT osteoporosis and fractures
 XX Claim 4; Page 122-123; 162pp; English.
 XX This shows a vervet transforming growth factor-beta (TGF-beta)
 CC binding protein designated VBEER. The cDNA and protein may be used for
 CC prevention, treatment and diagnosis of diseases associated with
 CC inappropriate BEER expression. For example, they may be used to treat
 CC disorders associated with decreased TGF-beta BP expression. The cDNA or
 CC vectors may be administered to treat diseases by rectifying mutations or
 CC deletions in a patient's genome that affect the activity of BEER by
 CC expressing inactive proteins or to supplement the patients own production
 CC of BEER polypeptides. The nucleic acids may be used for recombinant
 CC production of BEER, gene therapy, antisense therapy, as probes for
 CC diagnostic assays and for functional studies. BEER may be used to raise
 CC antibodies and for identification of BEER modulators. BEER antagonists
 CC may be used to increase bone mineral content for the treatment of
 CC disorders such as osteopenia, osteoporosis, fractures and other disorders
 CC associated with low mineral content.
 XX Sequence 213 AA;
 XX Query Match 17.3%; Score 32; DB 21; Length 213;
 XX Best Local Similarity 100.0%; Pred. No. 7.5e-25;
 XX Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX QY 65 TDGPCRSKAPVTELVCSCGCGPARLLPNAIGR 96
 XX |||||
 XX DB 90 TDGPCRSKAPVTELVCSCGCGPARLLPNAIGR 121
 XX
 XX RESULT 13
 XX AAY96436
 XX ID AAY96436 standard; Protein; 213 AA.
 XX AC AAY96436;
 XX DT 12-SEP-2000 (first entry)
 XX DE Human TGF-beta binding protein (BEER) variant P38R.
 XX osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
 XX BEER; variant; P38R; gene therapy; antisense therapy; fracture;
 XX chromosome 17q12-21; bone mineralization.
 XX Homo sapiens.
 XX Key Location/Qualifiers
 XX FT Misc-difference 38 /label= P38R
 XX FT /note= "wild type proline has been substituted with
 XX arginine"
 XX PN WO200032773-A1.
 XX 08-JUN-2000.
 XX 24-NOV-1999; 99WO-US27990.
 XX 27-NOV-1998; 98US-0110283.
 XX (DARW-) DARWIN DISCOVERY LTD.
 XX Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepers BW;
 XX Van Ness J, Winkler DG;
 XX WPI; 2000-412321/35.
 DR N-PSDB; AAA29057.
 XX Nucleic acids (I) encoding a transforming growth factor beta binding
 PT protein, useful for identifying agents for treating osteopenia,
 PT osteoporosis and fractures
 XX Claim 4; Page 122-123; 162pp; English.
 XX This shows a vervet transforming growth factor-beta (TGF-beta)
 CC binding protein designated VBEER. The cDNA and protein may be used for
 CC prevention, treatment and diagnosis of diseases associated with
 CC inappropriate BEER expression. For example, they may be used to treat
 CC disorders associated with decreased TGF-beta BP expression. The cDNA or
 CC vectors may be administered to treat diseases by rectifying mutations or
 CC deletions in a patient's genome that affect the activity of BEER by
 CC expressing inactive proteins or to supplement the patients own production
 CC of BEER polypeptides. The nucleic acids may be used for recombinant
 CC production of BEER, gene therapy, antisense therapy, as probes for
 CC diagnostic assays and for functional studies. BEER may be used to raise
 CC antibodies and for identification of BEER modulators. BEER antagonists
 CC may be used to increase bone mineral content for the treatment of
 CC disorders such as osteopenia, osteoporosis, fractures and other disorders
 CC associated with low mineral content.
 XX Sequence 213 AA;
 XX Query Match 17.3%; Score 32; DB 21; Length 213;
 XX Best Local Similarity 100.0%; Pred. No. 7.5e-25;
 XX Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX QY 65 TDGPCRSKAPVTELVCSCGCGPARLLPNAIGR 96
 XX |||||
 XX DB 90 TDGPCRSKAPVTELVCSCGCGPARLLPNAIGR 121
 XX
 XX RESULT 13
 XX AAY96436
 XX ID AAY96436 standard; Protein; 213 AA.
 XX AC AAY96436;
 XX DT 12-SEP-2000 (first entry)
 XX DE Human TGF-beta binding protein (BEER) variant P38R.
 XX osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
 XX BEER; variant; P38R; gene therapy; antisense therapy; fracture;
 XX chromosome 17q12-21; bone mineralization.
 XX Homo sapiens.
 XX Key Location/Qualifiers
 XX FT Misc-difference 38 /label= P38R
 XX FT /note= "wild type proline has been substituted with
 XX arginine"
 XX PN WO200032773-A1.
 XX 08-JUN-2000.
 XX 24-NOV-1999; 99WO-US27990.
 XX 27-NOV-1998; 98US-0110283.
 XX (DARW-) DARWIN DISCOVERY LTD.
 XX Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepers BW;
 XX Van Ness J, Winkler DG;

AAB26106
 ID AAB26106 standard; Protein; 213 AA.
 XX
 AC AAB26106;
 XX
 DT 15-JAN-2001 (first entry)
 XX
 DE Human DAN/Cerberus-related protein 6 (hDCR6) (exons 1 and 4).
 XX
 KW Human; DNA/Cerberus-related protein 6; hDCR6; morphogenic protein;
 KW antagonist; BMP; cell growth; cell differentiation; bone formation;
 KW gene therapy.
 XX
 OS Homo sapiens.
 XX
 PN WO200055193-A2.
 XX
 PD 21-SEP-2000.
 XX
 PF 02-MAR-2000; 2000WO-US05537.
 XX
 PR 12-MAR-1999; 99US-0124118.
 XX
 PA (REGG-) REGENERON PHARM INC.
 XX
 PI Economides AN;
 XX
 DR WPI; 2000-638179/61.
 DR N-PSDB; AAA94051.
 XX
 FT Novel isolated, human DNA/Cerberus related protein 6 which include
 FT natural homologue, and polypeptides comprising DCR6 domain and nucleic
 FT acids encoding the proteins which are useful as probes and primers -
 XX
 PS Claim 8; Fig 3; 40pp; English.
 XX
 CC The present sequence comprises the amino acid sequence encoded by exons 1
 CC and 4 of the human DAN/Cerberus-related protein 6 (hDCR6) coding
 CC sequence. The coding sequence was isolated from a human kidney cDNA
 CC library containing exons 1 and 4 of the sequence. hDCR6 is closely
 CC related to the DAN and DCR5 proteins, both of which act as antagonists of
 CC morphogenic proteins such as BMP. It is possible that the hDCR6 gene and
 CC protein can be used as immunogens, modulators of cell function, growth
 CC and differentiation, to reduce undesirable bone formation, to identify
 CC DCR6 binding agents, in diagnosis, and in gene therapy.
 XX
 SQ Sequence 213 AA;
 Query Match 17.3%; Score 32; DB 21; Length 213;
 Best Local Similarity 100.0%; Pred. No. 7.5e-25;
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 65 TDGFCRSKPVTELVCSGCGPARLLPNAIGR 96
 DB 90 TDGFCRSKPVTELVCSGCGPARLLPNAIGR 121
 RESULT 10
 AAY96429
 ID AAY96429 standard; Protein; 213 AA.
 XX
 AC AAY96429;
 XX
 DT 12-SEP-2000 (first entry)
 XX
 DE Human TGF-beta binding protein (BEER).
 DE osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
 KW BEER; gene therapy; antisense therapy; fracture; bone mineralization.
 KW
 OS Homo sapiens.
 XX
 PN WO200032773-A1.
 XX

XX 08-JUN-2000.
 XX
 XX 24-NOV-1999; 99WO-US27990.
 XX
 XX 27-NOV-1998; 98US-0110283.
 XX
 XX (DARW-) DARWIN DISCOVERY LTD.
 XX
 XX Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepers BW;
 PI Van Ness J, Winkler DG;
 XX
 DR WPI; 2000-412321/35.
 DR N-PSDB; AAA29055.
 XX
 FT Nucleic acids (I) encoding a transforming growth factor beta binding
 FT protein, useful for identifying agents for treating osteopenia,
 FT osteoporosis and fractures
 XX
 PS Claim 2; Page 116; 152pp; English.
 XX
 CC This shows the human transforming growth factor-beta (TGF-beta)
 CC binding protein designated hBEER. The cDNA and protein may be used for
 CC prevention, treatment and diagnosis of diseases associated with
 CC inappropriate BEER expression. For example, they may be used to treat
 CC disorders associated with decreased TGF-beta BP expression. The cDNA or
 CC vectors may be administered to treat diseases by rectifying mutations or
 CC deletions in a patient's genome that affect the activity of BEER by
 CC expressing inactive proteins or to supplement the patient's own production
 CC of BEER polypeptides. The nucleic acids may be used for recombinant
 CC production of BEER, gene therapy, antisense therapy, as probes for
 CC diagnostic assays and for functional studies. BEER may be used to raise
 CC antibodies and for identification of BEER modulators. BEER antagonists
 CC may be used to increase bone mineral content for the treatment of
 CC disorders such as osteopenia, osteoporosis, fractures and other
 CC disorders associated with low mineral content.
 XX
 SQ Sequence 213 AA;
 Query Match 17.3%; Score 32; DB 21; Length 213;
 Best Local Similarity 100.0%; Pred. No. 7.5e-25;
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 65 TDGFCRSKPVTELVCSGCGPARLLPNAIGR 96
 DB 90 TDGFCRSKPVTELVCSGCGPARLLPNAIGR 121
 RESULT 11
 AAY96430
 ID AAY96430 standard; Protein; 213 AA.
 XX
 AC AAY96430;
 XX
 DT 12-SEP-2000 (first entry)
 XX
 DE Human TGF-beta binding protein (BEER) variant V101.
 XX
 KW osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
 KW BEER; variant; V101; gene therapy; antisense therapy; fracture;
 KW chromosome 17q12-21; bone mineralization.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 10 /label= V101
 FT /note= "wild type valine has been substituted with
 FT isoleucine"
 XX
 PN WO200032773-A1.
 XX
 XX 08-JUN-2000.

RESULT 7

AAAY96434
ID AAAY96434 standard; Protein; 176 AA.

XX
AC AAAY96434;
XX
DT 12-SEP-2000 (first entry)

XX
DE Bovine TGF-beta binding protein (BEER).

XX
KW osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
KW BEER; gene therapy; antisense therapy; fracture; bone mineralization.

XX
OS Bos taurus.

XX
PN WO200032773-A1.

XX
PD 08-JUN-2000.

XX
PF 24-NOV-1999; 99WO-US27990.

XX
PR 27-NOV-1998; 98US-0110283.

XX
PA (DARW-) DARWIN DISCOVERY LTD.

XX
PI Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepers BW;
PI Van Ness J, Winkler DG;

XX
DR WPI; 2000-412321/35.

XX
DR N-PSDB; AAA29060.

XX
PT Nucleic acids (1) encoding a transforming growth factor beta binding
PT protein, useful for identifying agents for treating osteopenia,
PT osteoporosis and fractures

XX
PS Claim 7; Page 127; 162pp; English.

XX
CC This shows a bovine transforming growth factor-beta (TGF-beta)
CC binding protein designated bBEER. The cDNA and protein may be used for
CC prevention, treatment and diagnosis of diseases associated with
CC inappropriate BEER expression. For example, they may be used to treat
CC disorders associated with decreased TGF-beta Bp expression. The cDNA or
CC vectors may be administered to treat diseases by rectifying mutations or
CC deletions in a patient's genome that affect the activity of BEER by
CC expressing inactive proteins or to supplement the patients own production
CC of BEER polypeptides. The nucleic acids may be used for recombinant
CC production of BEER, gene therapy, antisense therapy, as probes for
CC diagnostic assays and for functional studies. BEER may be used to raise
CC antibodies and for identification of BEER modulators. BEER antagonists
CC may be used to increase bone mineral content for the treatment of
CC disorders such as osteopenia, osteoporosis, fractures and other disorders
CC associated with low mineral content.

XX
SQ Sequence 176 AA;

Query Match 17.3%; Score 32; DB 21; Length 176;
Best Local Similarity 100.0%; Pred. No. 6.4e-25;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 65 TDGPCRSAPVTELVCSGCGPARLLPNAIGR 96

DB 59 TDGPCRSAPVTELVCSGCGPARLLPNAIGR 90

RESULT 8

ABB07207
ID ABB07207 standard; Protein; 190 AA.

XX
AC ABB07207;

XX
DT 26-MAR-2002 (first entry)

XX

DE Human cloaked-2 polypeptide mature protein sequence.

XX
KW Cloaked-2; cysteine knot motif; nephrotropic; cardiant; immunomodulator;
KW hepatotropic; antiinflammatory; antithyroid; cyrostatic; neuroprotective;
KW antianemic; hypotensive; antiarrhythmic; antiarteriosclerotic; muscular;
KW antidiabetic; anorectic; gene therapy; cell therapy; antisense therapy;
KW human.

XX
OS Homo sapiens.

XX
FH Key Location/Qualifiers

FT Misc-difference 1.16

FT /note= "this region has been repeated twice in the
FT sequence provided in the sequence listing but has
FT been indicated correctly in the sequence in the
FT Figure"

XX
PN WO200192308-A2.

XX
PD 06-DEC-2001.

XX
PF 29-MAY-2001; 2001WO-US17478.

XX
PR 01-JUN-2000; 2000US-208550P.

XX
PR 04-AUG-2000; 2000US-223542P.

XX
PA (AMGE-) AMGEN INC.

XX
PI Paszty CJ, Gao Y;

XX
XX WPI; 2002-114325/15.

XX
DR N-PSDB; ABA94293.

XX
PT New human and mouse cysteine-knot polypeptide designated as Cloaked-2,
PT for treating or preventing kidney, heart (e.g. myocardial infarction)
PT or liver (e.g. hepatitis) diseases

XX
PS Claim 13; Fig 1; 170pp; English.

XX
CC The invention relates to polypeptides comprising a cysteine knot motif
CC and designated as Cloaked-2, derived from human and mouse. The cloaked-2
CC polypeptides can be expressed by standard recombinant methodology. The
CC cloaked-2 polynucleotides are useful in gene therapy and antisense
CC therapy. The cloaked-2 polypeptides and polynucleotides are useful for
CC treating, preventing, ameliorating or detecting diseases and disorders of
CC the kidney (e.g. anemia, hypertension or low blood pressure), heart (e.g.
CC cardiac hypertrophy, congestive heart failure, myocardial infarction,
CC arrhythmias, atherosclerosis, hypertension or low blood pressure),
CC skeletal muscle (e.g. muscular dystrophy or cachexia), placenta (e.g.
CC congenital abnormalities or miscarriage), liver (e.g. hepatitis or
CC cirrhosis), pancreas (e.g. diabetes or pancreatitis), thyroid (e.g.
CC Grave's disease or myxedema) or adrenal cortex (e.g. Cushing's disease
CC or Addison's disease), homeostasis or metabolic diseases (e.g. obesity,
CC cancer or myopathies), infections, or autoimmune diseases. Selective
CC binding agents may be used to modulate the biological activities of
CC Cloaked-2 polypeptides or to detect Cloaked-2 polypeptide levels in a
CC sample. Transgenic non-human animals are useful for drug candidate
CC screening. The present sequence represents the human cloaked-2
CC mature polypeptide sequence.

XX
SQ Sequence 190 AA;

Query Match 17.3%; Score 32; DB 23; Length 190;
Best Local Similarity 100.0%; Pred. No. 6.8e-25;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 65 TDGPCRSAPVTELVCSGCGPARLLPNAIGR 96

DB 67 TDGPCRSAPVTELVCSGCGPARLLPNAIGR 98

RESULT 9

XX Human; bone marrow expressed exon; gene expression analysis; probe;
 KW microarray; cancer; leukaemia; lymphoma; myeloma.
 XX Homo sapiens.
 XX WO200157276-A2.
 XX 09-AUG-2001.
 XX 30-JAN-2001; 2001WO-US00668.
 XX 04-FEB-2000; 2000US-0180312.
 XX 26-MAY-2000; 2000US-0207456.
 XX 30-JUN-2000; 2000US-0608408.
 XX 03-AUG-2000; 2000US-0632366.
 XX 21-SEP-2000; 2000US-0234687.
 XX 27-SEP-2000; 2000US-0236359.
 XX 04-OCT-2000; 2000GB-0024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-489900/53.
 XX Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human bone marrow -
 XX Example 4; SEQ ID NO: 33373; 689pp + Sequence Listing; English.
 XX The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC bone marrow. They can be used to measure gene expression in bone marrow
 CC samples, which may enable the improved diagnosis and treatment of cancers
 CC such as lymphoma, leukaemia and myeloma. The present sequence is a
 CC protein encoded by one of the probes of the invention.
 XX Sequence 139 AA;
 SQ
 Query Match 17.3%; Score 32; DB 22; Length 139;
 Best Local Similarity 100.0%; Pred. No. 5.2e-25;
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 65 TDGFCRSAPVTELVCSGCGPARLLPNAIGR 96
 Db 16 TDGFCRSAPVTELVCSGCGPARLLPNAIGR 47
 RESULT 6
 ID ABG42911 standard; Peptide; 139 AA.
 AC ABG42911;
 XX 19-AUG-2002 (first entry)
 XX Human peptide encoded by genome-derived single exon probe SEQ ID 32576.
 XX Human; single exon probe; asthma; lung cancer; COPD; ILD;
 KW chronic obstructive pulmonary disease; interstitial lung disease;
 KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
 KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
 KW Hermansky-Pudlak syndrome; Karagener's disease; pulmonary
 KW pulmonary histiocytosis; lymphangioleiomyomatosis; pulmonary
 KW pulmonary alveolar proteinosis; fibrocystic pulmonary syndrome;
 KW primary ciliary dyskinesia; pulmonary hypertension;
 KW hyaline membrane disease.
 XX Homo sapiens.
 XX WO200186003-A2.
 XX

PD 15-NOV-2001.
 XX 30-JAN-2001; 2001WO-US00665.
 XX 04-FEB-2000; 2000US-180312P.
 XX 26-MAY-2000; 2000US-207456P.
 XX 30-JUN-2000; 2000US-0608408.
 XX 03-AUG-2000; 2000US-0632366.
 XX 21-SEP-2000; 2000US-234687P.
 XX 27-SEP-2000; 2000US-236359P.
 XX 04-OCT-2000; 2000GB-0024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2002-114183/15.
 XX Spatially-addressable set of single exon nucleic acid probes, used to
 PT measure gene expression in human lung samples -
 XX Claim 27; SEQ ID No 32576; 634pp; English.
 XX The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human lung comprising single exon nucleic acid probes having one of
 CC 12614 nucleic acid sequences mentioned in the specification, or their
 CC complements or the 12387 open reading frames derived from the 12614
 CC probes. Also included are a microarray comprising the novel set of
 CC probes; the novel set of probes which hybridise at high stringency to a
 CC nucleic acid expressed in the human lung; measuring gene expression in a
 CC sample derived from human lung, comprising (a) contacting the array with
 CC a collection of detectably labeled nucleic acids derived from human lung
 CC mRNA, and (b) measuring the label detectably bound to each probe of
 CC the array; identifying exons in a eukaryotic genome, comprising
 CC (a) algorithmically predicting at least one exon from genomic sequences
 CC of the eukaryote; and (b) detecting specific hybridisation of detectably
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
 CC having a fragment identical to the predicted exon, the probe is included
 CC in the above mentioned microarray; assigning exons to a single gene,
 CC comprising (a) identifying exons from genomic sequence by the method
 CC above and (b) measuring the expression of each of the exons in several
 CC tissues and/or cell types using hybridisation to a single exon
 CC microarrays having a probe with the exon, where a common pattern of
 CC expression of the exons in the tissues and/or cell types indicates that
 CC the exons should be assigned to a single gene; a peptide comprising one
 CC of 12011 sequences, mentioned in the specification, or encoded by the
 CC probes/open reading frames (ORF). The probes are used for gene
 CC expression analysis, and for identifying exons in a gene, particularly
 CC using human lung derived mRNA and for the study of lung diseases
 CC such as asthma, lung cancer, chronic obstructive pulmonary disease
 CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
 CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
 CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
 CC haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
 CC pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic
 CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
 CC and hyaline membrane disease. The present sequence is a peptide/protein
 CC encoded by a single exon probe of the invention.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX Sequence 139 AA;
 SQ

Query Match 17.3%; Score 32; DB 23; Length 139;
 Best Local Similarity 100.0%; Pred. No. 5.2e-25;
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 65 TDGFCRSAPVTELVCSGCGPARLLPNAIGR 96
 Db 16 TDGFCRSAPVTELVCSGCGPARLLPNAIGR 47

AC	AA96433;
XX	
XX	12-SEP-2000 (first entry)
XX	
XX	Rat TGF-beta binding protein (BEER).
XX	osteopathic; transforming growth factor-beta; TGF-beta; binding protein; BEER; Gene therapy; antisense therapy; fracture; bone mineralization.
XX	
OS	Rattus norvegicus.
XX	
WO	WO200032773-A1.
XX	
PD	08-JUN-2000.
XX	
PF	24-NOV-1999; 99WO-US27990.
XX	
PR	27-NOV-1998; 98US-0110283.
XX	
PA	(DARW-) DARWIN DISCOVERY LTD.
XX	
PI	Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paeper BW;
PI	Van Ness J, Winkler DG;
XX	
DR	WPI; 2000-412321/35.
XX	N-PSDB; AAA29059.
XX	
PT	Nucleic acids (I) encoding a transforming growth factor beta binding protein, useful for identifying agents for treating osteopenia, osteoporosis and fractures
XX	
PS	Claim 6; Page 125-126; 162pp; English.
XX	This shows a rat transforming growth factor-beta (TGF-beta) binding protein designated rBEER. The cDNA and protein may be used for prevention, treatment and diagnosis of diseases associated with CC inappropriate BEER expression. For example, they may be used to treat disorders associated with decreased TGF-beta BP expression. The cDNA or vectors may be administered to treat diseases by rectifying mutations or deletions in a patient's genome that affect the activity of BEER by expressing inactive proteins or to supplement the patients own production of BEER polypeptides. The nucleic acids may be used for recombinant production of BEER, gene therapy, antisense therapy, as probes for diagnostic assays and for functional studies. BEER may be used to raise antibodies and for identification of BEER modulators. BEER antagonists may be used to increase bone mineral content for the treatment of disorders such as osteopenia, osteoporosis, fractures and other disorders associated with low mineral content.
SQ	Sequence 213 AA;
	Query Match 58.4%; Score 108; DB 21; Length 213;
	Best Local Similarity 100.0%; Pred.No. 9e-104;
	Matches 108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	65 TDGPCRSKAPVTVELVCSGCGPARLLPNAIGRVKWRNPGDFRCIPDRIYRAQRVQLLCP 124
Db	90 TDGPCRSKAPVTVELVCSGCGPARLLPNAIGRVKWRNPGDFRCIPDRIYRAQRVQLLCP 149
QY	125 GGAAPRSRRVLIVASCKKRLTRFHNQSELKDFGPETARPQGRKP RP 172
Db	150 GGAAPRSRRVLIVASCKKRLTRFHNQSELKDFGPETARPQGRKP RP 197
RESULT 5	
AM73067	
ID	AA73067 standard; Protein; 139 AA.
XX	
AC	AA73067;
XX	
DT	06-NOV-2001 (first entry)
DE	Human bone marrow expressed probe encoded protein SEQ ID NO: 33373.

PA (AMGE-) AMGEN INC.
 PI Paszty CJ, Gao Y;
 XX
 DR WPI; 2002-114325/15.
 DR N-PSDB; ABA94294.

XX New human and mouse cysteine-knot polypeptide designated as Cloaked-2,
 PT for treating or preventing kidney, heart (e.g. myocardial infarction)
 PT or liver (e.g. hepatitis) diseases

XX Claim 13; Fig 2; 170pp; English.

XX The invention relates to polypeptides comprising a cysteine knot motif
 CC and designated as Cloaked-2, derived from human and mouse. The cloaked-2
 CC polypeptides can be expressed by standard recombinant methodology. The
 CC cloaked-2 polynucleotides are useful in gene therapy and antisense
 CC therapy. The cloaked-2 polypeptides and polynucleotides are useful for
 CC treating, preventing, ameliorating or detecting diseases and disorders of
 CC the kidney (e.g. anemia, hypertension or low blood pressure), heart (e.g.
 CC cardiac hypertrophy, congestive heart failure, myocardial infarction,
 CC arrhythmias, atherosclerosis, hypertension or low blood pressure),
 CC skeletal muscle (e.g. muscular dystrophy or cachexia), placenta (e.g.
 CC congenital abnormalities or miscarriage), liver (e.g. hepatitis or
 CC cirrhosis), pancreas (e.g. diabetes or pancreatitis), thyroid (e.g.
 CC Grave's disease or myxedema) or adrenal cortex (e.g. Cushing's disease
 CC cancer or myopathies), infections, or autoimmune diseases. Selective
 CC binding agents may be used to modulate the biological activities of
 CC Cloaked-2 polypeptides or to detect Cloaked-2 polypeptide levels in a
 CC sample. Transgenic non-human animals are useful for drug candidate
 CC screening. The present sequence represents the mouse cloaked-2 mature
 CC polypeptide.

XX Sequence 188 AA;

Query Match 94.1%; Score 174; DB 23; Length 188;
 Best Local Similarity 100.0%; Pred. No. 2.4e-172;
 Matches 174; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGWAQFRNDATEVTPGLGEYPPPPNNQTMNRAENGRRPHHPYDAKDVSYSRELHY 60
 Db 1 QGWAQFRNDATEVTPGLGEYPPPPNNQTMNRAENGRRPHHPYDAKDVSYSRELHY 60
 QY 61 TRFLTGPCRSKAPVTELVCSCGCGPARLLPNAIGRVKWRPNPDPFCIPDRYRAQRVQ 120
 Db 61 TRFLTGPCRSKAPVTELVCSCGCGPARLLPNAIGRVKWRPNPDPFCIPDRYRAQRVQ 120
 QY 121 LLCPGAAPRSRKVRVLVASCCKKRLTRFNQSELKDFGPETARPOKGRKPRGA 174
 Db 121 LLCPGAAPRSRKVRVLVASCCKKRLTRFNQSELKDFGPETARPOKGRKPRGA 174

RESULT 2
 AAY96432

ID AAY96432 standard; Protein; 211 AA.

XX AAY96432;

XX 12-SEP-2000 (first entry)

DE Murine TGF-beta binding protein (BEER).

XX osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
 KW BEER; gene therapy; antisense therapy; fracture; bone mineralization.

XX Mus musculus.

XX WO200032773-A1.

XX 08-JUN-2000.

XX 24-NOV-1999; 99WO-US27990.

XX 27-NOV-1998; 98US-0110283.
 XX (DAEW-) DARWIN DISCOVERY LTD.

XX Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepers BW;
 PI Van Ness J, Winkler DG;
 XX WPI; 2000-412321/35.
 DR N-PSDB; AAA29058.

XX Nucleic acids (1) encoding a transforming growth factor beta binding
 PT protein, useful for identifying agents for treating osteopenia,
 PT osteoporosis and fractures

XX Claim 5; Page 124; 152pp; English.

XX This shows a murine transforming growth factor-beta (TGF-beta)
 CC binding protein designated mBEER. The cDNA and protein may be used for
 CC prevention, treatment and diagnosis of diseases associated with
 CC inappropriate BEER expression. For example, they may be used to treat
 CC disorders associated with decreased TGF-beta BP expression. The cDNA or
 CC vectors may be administered to treat diseases by rectifying mutations or
 CC deletions in a patient's genome that affect the activity of BEER by
 CC expressing inactive proteins or to supplement the patients own production
 CC of BEER polypeptides. The nucleic acids may be used for recombinant
 CC production of BEER, gene therapy, antisense therapy, as probes for
 CC diagnostic assays and for functional studies. BEER may be used to raise
 CC antibodies and for identification of BEER modulators. BEER antagonists
 CC may be used to increase bone mineral content for the treatment of
 CC disorders such as osteopenia, osteoporosis, fractures and other disorders
 CC associated with low mineral content.

XX Sequence 211 AA;

Query Match 94.1%; Score 174; DB 21; Length 211;
 Best Local Similarity 100.0%; Pred. No. 2.6e-172;
 Matches 174; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGWAQFRNDATEVTPGLGEYPPPPNNQTMNRAENGRRPHHPYDAKDVSYSRELHY 60
 Db 24 QGWAQFRNDATEVTPGLGEYPPPPNNQTMNRAENGRRPHHPYDAKDVSYSRELHY 83
 QY 61 TRFLTGPCRSKAPVTELVCSCGCGPARLLPNAIGRVKWRPNPDPFCIPDRYRAQRVQ 120
 Db 84 TRFLTGPCRSKAPVTELVCSCGCGPARLLPNAIGRVKWRPNPDPFCIPDRYRAQRVQ 143
 QY 121 LLCPGAAPRSRKVRVLVASCCKKRLTRFNQSELKDFGPETARPOKGRKPRGA 174
 Db 144 LLCPGAAPRSRKVRVLVASCCKKRLTRFNQSELKDFGPETARPOKGRKPRGA 197

RESULT 3
 ABB07210

ID ABB07210 standard; Protein; 211 AA.

XX ABB07210;

XX 26-MAR-2002 (first entry)

DE Mouse cloaked-2 polypeptide sequence.

XX Cloaked-2; cysteine knot motif; nephrotropic; cardiant; immunomodulator;
 KW hepatotropic; antiinflammatory; antithyroid; cytostatic; neuroprotective;
 KW antianemic; hypotensive; antiarrhythmic; antiarteriosclerotic; muscular;
 KW antidiabetic; anorectic; gene therapy; cell therapy; antisense therapy;
 KW mouse.

XX Mus musculus.

XX Key

XX Location/Qualifiers

XX Peptide

XX 1..23
 /note= "signal peptide"

Result No.	Score	Query Match	Length	DB	ID	Description
1	174	94.1	188	23	AB07208	Mouse cloaked-2 po
2	174	94.1	211	21	AAV96432	Mouse TGF-beta bi
3	174	94.1	211	23	AB07210	Mouse cloaked-2 po
4	108	58.4	213	21	AAV96433	Rat TGF-beta bind1
5	32	17.3	139	22	AAW73067	Human bone marrow
6	32	17.3	139	23	ABG42911	Human peptide enco
7	32	17.3	176	21	AAV96434	Bovine TGF-beta bi
8	32	17.3	190	23	AB07207	Human cloaked-2 po
9	32	17.3	213	21	AAB26106	Human DAN/Cerberus
10	32	17.3	213	21	AAV96429	Human TGF-beta bin

SUMMARIES

PR 04-AUG-2000; 200003-223344: yy

PR 04-AUG-2000; 200003-223344: yy

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA: US/10/159,749
FILING DATE: 29-May-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/040,229
FILING DATE: 13-Mar-1998
ATTORNEY/AGENT INFORMATION:
NAME: OSMAN, RICHARD A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: B97-020-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 343-4341
TELEFAX: (650) 343-4342
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 184 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-10-159-749-6

Query Match 8.8%; Score 90.5; DB 9; Length 184;
Best Local Similarity 22.9%; Pred. No. 0.21; Mismatches 60; Indels 35; Gaps 8;
Matches 35; Conservative 23;

QY 18 GEYEP-----PPENNTMNAENGRRPHH-----PYDAKDVSEYSCRELHYT--RFLTDG 67
Db 33 GAIPPPDKQPNDSQWQTOQSQSHRERGKGTSMPEAEVLESSQEAHLHITERKYLKED 92
QY 68 PCR-----SAKPVTELCSGCGPARLLPNAIGRVKWRPNPDPFR-----CI 110
Db 93 WCKTQPLKQTIHEEGCNSRTIINRFYGCQN-SFYIPRHV-----RKEGSGFQSCSFK 145
QY 111 PDYRAQRVQLLCPGGAAPRSRK-VRLVASCCK 142
Db 146 PKKFTTMTVNLNCPQLQPPRKXKRITRVKQCR 178

RESULT 14
US-10-159-749-4
Sequence 4, Application US/10159749
Publication No. US20020192219A1
GENERAL INFORMATION:
APPLICANT: Harland, Richard
Hsu, David
TITLE OF INVENTION: Morphogenic Proteins
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 75 DENISE DRIVE
CITY: HILLSBOROUGH
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94010
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/159,749
FILING DATE: 29-May-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/040,229
FILING DATE: 13-Mar-1998
ATTORNEY/AGENT INFORMATION:
NAME: OSMAN, RICHARD A
REGISTRATION NUMBER: 36,627

REFERENCE/DOCKET NUMBER: B97-020-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 343-4341
TELEFAX: (650) 343-4342
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 182 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-159-749-4

Query Match 8.6%; Score 88.5; DB 9; Length 182;
Best Local Similarity 23.5%; Pred. No. 0.32; Mismatches 54; Indels 37; Gaps 8;
Matches 35; Conservative 23;

QY 23 PPENNTMNAENG-----GRPPHHYPYDAKDVSEYSCRELHYT--RFLTDGPCR- 70
Db 36 PPDKGQP-NDSEOGQAPGDRVRGKGGQALAAEEVLESSQEAHLHITERKYLKRDWCKT 94
QY 71 -----SAKPVTELCSGCGPARLLPNAIGRVKWRPNPDPFR-----CIPDRY 114
Db 95 QPLKQTIHEEGCNSRTIINRFYGCQN-SFYIPRHIRE-----GSFQSCSFCCKPK 147
QY 115 RAQRVQLLCPGGAAP-RSRKRVLVASCCK 142
Db 148 TTMVTLNCPQLQPPRKXKRITRVKQCR 176

RESULT 15
US-10-044-716-10
Sequence 10, Application US/10044716
Patent No. US20020159986A1
GENERAL INFORMATION:
APPLICANT: LANGENFELD, John
TITLE OF INVENTION: BONE MORPHOGENETIC PROTEIN-2 IN THE TREATMENT AND DIAGNOSIS OF CANCER
FILE REFERENCE: 270/070US
CURRENT APPLICATION NUMBER: US/10/044,716
CURRENT FILING DATE: 2002-01-11
PRIOR APPLICATION NUMBER: US60/261,252
PRIOR FILING DATE: 2001-01-12
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patent in version 3.1
SEQ ID NO 10
TYPE: PRT
LENGTH: 184
ORGANISM: Homo sapiens
US-10-044-716-10

Query Match 8.6%; Score 88.5; DB 9; Length 184;
Best Local Similarity 23.4%; Pred. No. 0.32; Mismatches 55; Indels 43; Gaps 8;
Matches 36; Conservative 20;

QY 23 PPENNTMNAENG-----GRPPHHYPYDAKDVSEYSCRELHYT--RFLTDG 66
Db 34 PPDKAQNDSEQTQSPQPGSRNRGRGQGTAMP--GEEVLESSQEAHLHITERKYLK 91
QY 67 GPCR-----SAKPVTELCSGCGPARLLPNAIGRVKWRPNPDPFR-----C 109
Db 92 DWCKTQPLKQTIHEEGCNSRTIINRFYGCQN-SFYIPRHIRE-----RKEGSGFQSCSFC 144
QY 110 IPDRAQRVQLLCPGGAAP-RSRKRVLVASCCK 142
Db 145 KKKFTTMTVNLNCPQLQPPRKXKRITRVKQCR 178

Search completed: March 28, 2003, 14:20:45
Job time : 11.84 secs

```

Db 159 LGAEFDGLGLPTPERSLDGL-----EAGEAFFFFPPAAPTARWRPSAP-----PTAHRV 209
QY 117 -----QVQLLCPGA-----APRSKRVLRVAS-----CKKRLTRFHNSQLKDFGP 159
Db 210 VAGPRRLRSGGGGQADRAPRARSRACTTAPWARPARTPARCRR----- 254
QY 160 ETARPOKGRKRRPGAK 175
Db 255 -RPRPRWRAPSPGAR 269

RESULT 12
US-09-814-777A-99
; Sequence 99, Application US/09814777A
; Patent No. US20020142415A1
; GENERAL INFORMATION:
; APPLICANT: KOOPMAN, Peter Anthony
; APPLICANT: MUSCAT, George Eugene Orlando
; TITLE OF INVENTION: NOVEL POLYPEPTIDES AND POLYNUCLEOTIDES AND METHODS OF USING THEM
; FILE REFERENCE: 21415-0003
; CURRENT APPLICATION NUMBER: US/09/814,777A
; PRIOR FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: AU PQ6457
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 128
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 99
; LENGTH: 462
; TYPE: PRT
; ORGANISM: Human
US-09-814-777A-99

Query Match 9.3%; Score 95.5; DB 10; Length 462;
Best Local Similarity 23.1%; Pred. No. 0.21;
Matches 45; Conservative 23; Mismatches 64; Indels 63; Gaps 10;

QY 1 QGWAQRNDATFVIGLGEYPEPPENNQTWNAENG--GRPPHPHYDAKOVSEYSREL 58
Db 117 QARKARRLEPGLLLPGLAP-PQPPPEPPAASGSAFRGSPAGAEFDGLGL----- 167
QY 59 HYTRFLTDGCRSAKPVTELVCSGCGPARLL-PNAIGRVKWRPNGPDPFRCPDRYRA- 116
Db 168 -----PTPERSLDGL-----EAGEAFFFFPPAAPTARWRPSAP-----PTAHRV 209
QY 117 ---QVQLLCPGA-----APRSKRVLRVAS-----CKKRLTRFHNSQLKDFGPE 160
Db 210 AGPRRLRSGGGGQADRAPRARSRACTTAPWARPARTPARCRR----- 253
QY 161 TARPOKGRKRRPGAK 175
Db 254 RPRPRWRAPSPGAR 269

RESULT 13
US-10-159-749-6
; Sequence 6, Application US/10159749
; Publication No. US20020192219A1
; GENERAL INFORMATION:
; APPLICANT: Harland, Richard
; Hsu, David
; TITLE OF INVENTION: Morphogenic Proteins
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 75 DENISE DRIVE
; CITY: HILLSBOROUGH
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94010
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```



```

GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; TITLE OF INVENTION: and Methods for Their Use
; FILE REFERENCE: 11000.101LC4U
; CURRENT APPLICATION NUMBER: US/09/866,050A
; CURRENT FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 725
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 159
; LENGTH: 206
; TYPE: PRT
; ORGANISM: Mouse
US-09-866-050A-159

Query Match      32.0%; Score 330; DB 9; Length 206;
Best Local Similarity 39.8%; Pred. No. 1.6e-24;
Matches 74; Conservative 35; Mismatches 65; Indels 12; Gaps 7;

QY    5 AFRNDATETVIGLGEYPPEP-PENNQTWNAENGCRPPHPH---YDAKDVSSEYSCELHY 60
      ||:|||::||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db    23 AFKNDATEILYSHVKVPFAHPSSNSTLNQARNGR--HFSTGLDRNSRVQGCELEAS 80

QY    61 TRFTLDGPCRSAPKVTELVCSGQCQPALLPNIAIG--RVKWM-RPNGPDPRICIPDRYRA 116
      ||:|||::||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db    81 TKIISDGQCTSIPLKELVCAGECLPLVLPNWGGYGTYKWSRSSQEWRGVNDKTRET 140

QY    117 QRVLQLCPGGAAPSRRKVRIVASCKKLTRHNOSSELKDFGPETARP-OKGRKRPFPAK 175
      ||:|||::||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db    141 QRIQLCQODGST-RYKITVTACKCRYTROHNESSHNFESVSPAKPAQHHRERKASK 199

QY    176 ANOAEI 181
      ::|||::|
Db    200 SSKHSL 205


RESULT 7
US-09-866-050A-286
; Sequence 286, Application US/09866050A
; Publication No. US2003004071A1
GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; TITLE OF INVENTION: and Methods for Their Use
; FILE REFERENCE: 11000.101LC4U
; CURRENT APPLICATION NUMBER: US/09/866,050A
; CURRENT FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 725
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 286
; LENGTH: 206
; TYPE: PRT
; ORGANISM: Mouse
US-09-866-050A-286

Query Match      32.0%; Score 330; DB 9; Length 206;
Best Local Similarity 39.8%; Pred. No. 1.6e-24;
Matches 74; Conservative 35; Mismatches 65; Indels 12; Gaps 7;

QY    5 AFRNDATETVIGLGEYPPEP-PENNQTWNAENGCRPPHPH---YDAKDVSSEYSCELHY 60
      ||:|||::||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db    23 AFKNDATEILYSHVKVPFAHPSSNSTLNQARNGR--HFSTGLDRNSRVQGCELEAS 80
```

RESULT 2
US-09-867-274-6
; Sequence 6, Application US/09867274
; Patent No. US20020106650A1
; GENERAL INFORMATION:
; APPLICANT: Paszty, Christopher
; APPLICANT: Gao, Yongming
; TITLE OF INVENTION: Cysteine Knot Polypeptides: Cloaked-2 Molecules and Uses Thereof
; FILE REFERENCE: 01017/37428
; CURRENT APPLICATION NUMBER: US/09/867,274
; PRIOR FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: US 60/208,550
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: US 60/223,542
; PRIOR FILING DATE: 2000-08-04
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 208
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-867-274-6

Query Match 100.0%; Score 1032; DB 10; Length 208;
Best Local Similarity 100.0%; Pred. No. 2e-92;
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGWAQFRNDATVIFGLGEYEPPEPP--ENNQTMNRAENGRRPHHPYDAKDVSEYSCREL 60
DB 24 QGWAQFRNDATVIFGLGEYEPPEPP--ENNQTMNRAENGRRPHHPYDAKDVSEYSCREL 83
QY 61 TRFLTDGFCRSAPKPVTELVCSCGCGPARLLPNAIGRVKWWRPNGPDRYRAQRVQ 120
DB 84 TRFLTDGFCRSAPKPVTELVCSCGCGPARLLPNAIGRVKWWRPNGPDRYRAQRVQ 143
QY 121 LLCPGGAAPRRKRVLVASCKCKRLTRFHNSQELKDFGTEAARPKQGRKPRPARSAK 180
DB 144 LLCPGGAAPRRKRVLVASCKCKRLTRFHNSQELKDFGTEAARPKQGRKPRPARSAK 203
QY 181 LENAY 185
DB 204 LENAY 208

RESULT 3
US-09-867-274-2
; Sequence 2, Application US/09867274
; Patent No. US20020106650A1
; GENERAL INFORMATION:
; APPLICANT: Paszty, Christopher
; APPLICANT: Gao, Yongming
; TITLE OF INVENTION: Cysteine Knot Polypeptides: Cloaked-2 Molecules and Uses Thereof
; FILE REFERENCE: 01017/37428
; CURRENT APPLICATION NUMBER: US/09/867,274
; CURRENT FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: US 60/208,550
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: US 60/223,542
; PRIOR FILING DATE: 2000-08-04
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 190
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-867-274-2

Query Match 90.4%; Score 932.5; DB 10; Length 190;
Best Local Similarity 88.4%; Pred. No. 7.4e-83;
Matches 168; Conservative 10; Mismatches 7; Indels 5; Gaps 2;

QY 1 QGWAQFRNDATVIFGLGEYEPPEPP--ENNQTMNRAENGRRPHHPYDAKDVSEYSCREL 58
DB 24 QGWAQFRNDATVIFGLGEYEPPEPP--ENNQTMNRAENGRRPHHPYDAKDVSEYSCREL 60
QY 59 HYTRFLTDGFCRSAPKPVTELVCSCGCGPARLLPNAIGRVKWWRPNGPDRYRAQR 118
DB 61 HFTRYVTDGFCRSAPKPVTELVCSCGCGPARLLPNAIGRVKWWRPNGPDRYRAQR 120
QY 119 VOLLCPGGAAPRRKRVLVASCKCKRLTRFHNSQELKDFGTEAARPKQGRKPRPARSAK 175
DB 121 VOLLCPGGAAPRRKRVLVASCKCKRLTRFHNSQELKDFGTEAARPKQGRKPRPARSAK 180
QY 176 ANQAELENAY 185
DB 181 ANQAELENAY 190

RESULT 4
US-09-867-274-5
; Sequence 5, Application US/09867274
; Patent No. US20020106650A1
; GENERAL INFORMATION:
; APPLICANT: Paszty, Christopher
; APPLICANT: Gao, Yongming
; TITLE OF INVENTION: Cysteine Knot Polypeptides: Cloaked-2 Molecules and Uses Thereof
; FILE REFERENCE: 01017/37428
; CURRENT APPLICATION NUMBER: US/09/867,274
; CURRENT FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: US 60/208,550
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: US 60/223,542
; PRIOR FILING DATE: 2000-08-04
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 213
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-867-274-5

Query Match 90.4%; Score 932.5; DB 10; Length 213;
Best Local Similarity 88.4%; Pred. No. 8.5e-83;
Matches 168; Conservative 10; Mismatches 7; Indels 5; Gaps 2;

QY 1 QGWAQFRNDATVIFGLGEYEPPEPP--ENNQTMNRAENGRRPHHPYDAKDVSEYSCREL 58
DB 24 QGWAQFRNDATVIFGLGEYEPPEPP--ENNQTMNRAENGRRPHHPYDAKDVSEYSCREL 83
QY 59 HYTRFLTDGFCRSAPKPVTELVCSCGCGPARLLPNAIGRVKWWRPNGPDRYRAQR 118
DB 84 HFTRYVTDGFCRSAPKPVTELVCSCGCGPARLLPNAIGRVKWWRPNGPDRYRAQR 143
QY 119 VOLLCPGGAAPRRKRVLVASCKCKRLTRFHNSQELKDFGTEAARPKQGRKPRPARSAK 175
DB 144 VOLLCPGGAAPRRKRVLVASCKCKRLTRFHNSQELKDFGTEAARPKQGRKPRPARSAK 203
QY 176 ANQAELENAY 185
DB 204 ANQAELENAY 213

RESULT 5
US-09-864-761-47109
; Sequence 47109, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wenheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1

GenCore version 5.1.4 p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 28, 2003, 14:16:22 ; Search time 11.84 Seconds
(without alignments)
917.557 Million cell updates/sec

Title: US-09-867-274-4
Perfect score: 1032
Sequence: 1 CQWQAFRNDATVIFGLGEY.....KGRKPRGAKANQALENAY 185

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 237916 seqs, 58723674 residues
Total number of hits satisfying chosen parameters: 237916

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : Published Applications AA:*
- 1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
 - 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
 - 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
 - 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
 - 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
 - 6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
 - 7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
 - 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
 - 9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
 - 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
 - 11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
 - 12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
 - 13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
 - 14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1032	100.0	185	10	US-09-867-274-4
2	1032	100.0	208	10	US-09-867-274-6
3	932.5	90.4	190	10	US-09-867-274-2
4	932.5	90.4	213	10	US-09-867-274-5
5	687.5	66.6	139	10	US-09-864-761-47109
6	330	32.0	206	9	US-09-866-050A-159
7	330	32.0	206	9	US-09-866-050A-286
8	328	31.8	206	10	US-09-853-625B-2
9	328	31.8	206	10	US-09-853-625B-20
10	324	31.4	183	10	US-09-867-274-25
11	110	10.7	337	10	US-09-814-777A-126
12	95.5	9.3	462	10	US-09-814-777A-99
13	90.5	8.8	184	9	US-10-159-749-6
14	88.5	8.6	182	9	US-10-159-749-4
15	88.5	8.6	184	9	US-10-044-716-10
16	88.5	8.6	184	9	US-10-159-749-2
17	86	8.3	147	9	US-10-159-749-9
18	83.5	8.1	147	9	US-10-092-154-828
19	83.5	8.1	147	10	US-09-764-847-828

20	80.5	7.8	270	9	US-09-903-170C-1
21	80.5	7.8	270	10	US-09-903-180B-1
22	80.5	7.8	270	10	US-09-903-187A-1
23	80.5	7.8	270	10	US-09-903-171A-1
24	80.5	7.8	270	10	US-09-903-188A-1
25	80.5	7.8	270	10	US-09-903-323A-1
26	80.5	7.8	270	10	US-09-903-325A-1
27	80	7.8	188	10	US-09-349-954A-18
28	80	7.8	188	10	US-09-907-007-18
29	80	7.8	251	10	US-09-764-853-522
30	80	7.8	251	10	US-09-764-898-176
31	76.5	7.4	141	9	US-09-813-398-3
32	76.5	7.4	145	9	US-09-760-294-12
33	76.5	7.4	145	9	US-09-760-294-13
34	76.5	7.4	145	9	US-09-760-294-14
35	76.5	7.4	145	9	US-09-760-294-15
36	76.5	7.4	145	9	US-10-131-241-65
37	76.5	7.4	165	9	US-09-915-676-1
38	76.5	7.4	165	10	US-09-466-320-14
39	76.5	7.4	267	9	US-10-044-716-12
40	76.5	7.4	267	9	US-09-089-818B-8
41	76.5	7.4	270	9	US-10-072-349-145
42	76.5	7.4	270	10	US-09-764-855-145
43	76.5	7.4	307	10	US-09-756-186-4
44	76.5	7.4	336	10	US-09-756-186-8
45	76.5	7.4	432	12	US-10-062-254-242

ALIGNMENTS

RESULT 1
US-09-867-274-4
; Sequence 4, Application US/09867274
; Patent No. US20020106650A1
; GENERAL INFORMATION:
; APPLICANT: Paszty, Christopher
; APPLICANT: Gao, Yongming
; TITLE OF INVENTION: Cysteine Knot Polypeptides: Cloaked-2 Molecules and Uses Thereof
; FILE REFERENCE: 01017/37428
; CURRENT APPLICATION NUMBER: US/09/867,274
; CURRENT FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: US 60/208,550
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: US 60/223,542
; PRIOR FILING DATE: 2000-08-04
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 185
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-867-274-4

Query Match	100.0%	Score	1032	DB	10	Length	185
Best Local Similarity	100.0%	Pred. No.	1.7e-92				
Matches	185	Conservative	0	Mismatches	0	Indels	0
Gaps	0						
QY	1	CQWQAFRNDATVIFGLGEYPPPPENNQTWNAENGRRPHHPYDAKVDKVSCEHLHY	60				
Db	1	CQWQAFRNDATVIFGLGEYPPPPENNQTWNAENGRRPHHPYDAKVDKVSCEHLHY	60				
QY	61	TFPLFDGFCRAKPVTELVCSSGCGPARLLPNAIGRVKWRPNGPDPFCIPDRYRAQRVQ	120				
Db	61	TFPLFDGFCRAKPVTELVCSSGCGPARLLPNAIGRVKWRPNGPDPFCIPDRYRAQRVQ	120				
QY	121	LLCPGGAAPRSRKVRLVASCKKCLTRFHNQSELKDFGPETARPOKGRKPRPGAKANQAE	180				
Db	121	LLCPGGAAPRSRKVRLVASCKKCLTRFHNQSELKDFGPETARPOKGRKPRPGAKANQAE	180				
QY	181	LENAY 185					
Db	181	LENAY 185					

Sequence 1, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 18, Appli
Sequence 18, Appli
Sequence 522, App
Sequence 176, App
Sequence 3, Appli
Sequence 12, Appli
Sequence 13, Appli
Sequence 14, Appli
Sequence 65, Appli
Sequence 1, Appli
Sequence 14, Appli
Sequence 12, Appli
Sequence 8, Appli
Sequence 145, App
Sequence 145, App
Sequence 4, Appli
Sequence 8, Appli
Sequence 242, App

; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 3814
; LENGTH: 198285
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 X87344
US-09-880-107-3814

Query Match 3.0%; Score 19; DB 10; Length 198285;
Best Local Similarity 100.0%; Pred. No. 1;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 AAGCCTTCAGGAATGTC 98
|||
DB 70568 AAGCCTTCAGGAATGTC 70586

RESULT 15

US-09-867-550-675
; Sequence 675, Application US/09867550
; Patent No. US20020082206A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Mehraban, Fuad,
; APPLICANT: Conley, Pamela
; APPLICANT: Law, Debbie
; APPLICANT: Topper, James
; TITLE OF INVENTION: No. US20020082206A1 Polynucleotides from Atherogenic Cells and
; FILE REFERENCE: 21402-013 (Cura-313)
; CURRENT APPLICATION NUMBER: US/09/867,550
; CURRENT FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: USSN 60/208,427
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 2125
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 675
; LENGTH: 323
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-867-550-675

Query Match 2.8%; Score 18; DB 10; Length 323;
Best Local Similarity 100.0%; Pred. No. 7.8;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 545 CGGAGACCGCGCGCGC 562
|||
DB 146 CGGAGACCGCGCGCGC 163

Search completed: March 29, 2003, 10:50:10
Job time : 115.594 secs

; PRIOR APPLICATION NUMBER: US 60/223,542
; PRIOR FILING DATE: 2000-08-04
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 29
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial: PCR primer
US-09-867-274-9

Query Match 3.3%; Score 21; DB 10; Length 29;
Best Local Similarity 100.0%; Pred. No. 0.27;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 68 GCCAGGGGTGGCAAGCTTCA 88
Db 1 GCCAGGGGTGGCAAGCTTCA 21
|||||

RESULT 12

US-09-864-761-9695/c
; Sequence 9695, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wenheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeonica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1

; SEQ ID NO 9695
; LENGTH: 577
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC002472.3
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 15
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 47
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 17
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 59
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 25
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 22
US-09-864-761-9695

Query Match 3.3%; Score 21; DB 10; Length 577;
Best Local Similarity 100.0%; Pred. No. 0.19;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 58 GCTGTGGAGGGCCAGGGGTGG 78
Db 474 GCTGTGGAGGGCCAGGGGTGG 454
|||||

RESULT 13

US-09-867-274-10/c
; Sequence 10, Application US/09867274
; Patent No. US20020106650A1
; GENERAL INFORMATION:
; APPLICANT: Paszty, Christopher
; APPLICANT: Gao, Yongming
; TITLE OF INVENTION: Cysteine Knot Polypeptides: Cloaked-2 Molecules and Uses Thereof
; FILE REFERENCE: 01017/37428
; CURRENT APPLICATION NUMBER: US/09/867,274
; PRIOR FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: US 60/208,550
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: US 60/223,542
; PRIOR FILING DATE: 2000-08-04
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 10
; LENGTH: 24
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial: PCR primer
US-09-867-274-10

Query Match 3.1%; Score 20; DB 10; Length 24;
Best Local Similarity 100.0%; Pred. No. 0.95;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 388 TTCGCTGCATCCCGGATCG 407
Db 20 TTCGCTGCATCCCGGATCG 1
|||||

RESULT 14

US-09-880-107-3814
; Sequence 3814, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379

Patent No. US20020106650A1
; GENERAL INFORMATION:
; APPLICANT: Paszty, Christopher
; TITLE OF INVENTION: Cysteine Knot Polypeptides: Cloaked-2 Molecules and Uses Thereof
; FILE REFERENCE: 01017/37428
; CURRENT APPLICATION NUMBER: US/09/867,274
; CURRENT FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: US 60/208,550
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: US 60/223,542
; PRIOR FILING DATE: 2000-08-04
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 19
; LENGTH: 45
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial: PCR primer
US-09-867-274-19

Query Match 3.9%; Score 25; DB 10; Length 45;
Best Local Similarity 100.0%; Pred. No. 0.002;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGCAGCCCTCACTAGCCCGTGCC 25
Db 21 ATGCAGCCCTCACTAGCCCGTGCC 45

RESULT 8
US-09-867-274-12/c
; Sequence 12, Application US/09867274
; Patent No. US20020106650A1
; GENERAL INFORMATION:
; APPLICANT: Paszty, Christopher
; TITLE OF INVENTION: Cysteine Knot Polypeptides: Cloaked-2 Molecules and Uses Thereof
; FILE REFERENCE: 01017/37428
; CURRENT APPLICATION NUMBER: US/09/867,274
; CURRENT FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: US 60/208,550
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: US 60/223,542
; PRIOR FILING DATE: 2000-08-04
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 24
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial: PCR primer
US-09-867-274-12

Query Match 3.8%; Score 24; DB 10; Length 24;
Best Local Similarity 100.0%; Pred. No. 0.0072;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 241 CTGCACCTACACCGCTTCTCTGACA 264
Db 24 CTGCACCTACACCGCTTCTCTGACA 1

RESULT 9
US-09-867-274-15
; Sequence 15, Application US/09867274
; Patent No. US20020106650A1
; GENERAL INFORMATION:
; APPLICANT: Paszty, Christopher
; TITLE OF INVENTION: Cysteine Knot Polypeptides: Cloaked-2 Molecules and Uses Thereof

FILE REFERENCE: 01017/37428
; CURRENT APPLICATION NUMBER: US/09/867,274
; CURRENT FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: US 60/208,550
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: US 60/223,542
; PRIOR FILING DATE: 2000-08-04
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 15
; LENGTH: 21
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial: PCR primer
US-09-867-274-15

Query Match 3.3%; Score 21; DB 10; Length 21;
Best Local Similarity 100.0%; Pred. No. 0.28;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 247 TACACCGCTTCTCTGACAGAC 267
Db 1 TACACCGCTTCTCTGACAGAC 21

RESULT 10
US-09-867-274-17
; Sequence 17, Application US/09867274
; Patent No. US20020106650A1
; GENERAL INFORMATION:
; APPLICANT: Paszty, Christopher
; TITLE OF INVENTION: Cysteine Knot Polypeptides: Cloaked-2 Molecules and Uses Thereof
; FILE REFERENCE: 01017/37428
; CURRENT APPLICATION NUMBER: US/09/867,274
; CURRENT FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: US 60/208,550
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: US 60/223,542
; PRIOR FILING DATE: 2000-08-04
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 17
; LENGTH: 21
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial: PCR primer
US-09-867-274-17

Query Match 3.3%; Score 21; DB 10; Length 21;
Best Local Similarity 100.0%; Pred. No. 0.28;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 291 GGTCACCGAGTTGGTG*GCTC 311
Db 1 GGTCACCGAGTTGGTG*GCTC 21

RESULT 11
US-09-867-274-9
; Sequence 9, Application US/09867274
; Patent No. US20020106650A1
; GENERAL INFORMATION:
; APPLICANT: Paszty, Christopher
; TITLE OF INVENTION: Cysteine Knot Polypeptides: Cloaked-2 Molecules and Uses Thereof
; FILE REFERENCE: 01017/37428
; CURRENT APPLICATION NUMBER: US/09/867,274
; CURRENT FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: US 60/208,550
; PRIOR FILING DATE: 2000-06-01

; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 3098
; LENGTH: 422
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC003098.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.62
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.6
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.49
; OTHER INFORMATION: EST HUMAN HIT: BE613498.1, EVALUATE 9.90e-02
; OTHER INFORMATION: SWISSPROT HIT: P45646, EVALUATE 4.90e-01
; OTHER INFORMATION: NT HIT: AF074705.1, EVALUATE 1.00e+00
US-09-864-761-30988

Query Match 10.8%; Score 69; DB 10; Length 422;
Best Local Similarity 100.0%; Pred. No. 7.5e-27;
Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 475 CTGGTGGCTCGTGAAGTGAAGCGCCTCACCGCTTCCACCAACGAGTCGAGCTCAAG 534
DB 162 CTGGTGGCTCGTGAAGTGAAGCGCCTCACCGCTTCCACCAACGAGTCGAGCTCAAG 103
QY 535 GACTTCGGG 543
DB 102 GACTTCGGG 94

RESULT 4
US-09-867-274-1
; Sequence 1, Application US/09867274
; Patent No. US20020106650A1
; GENERAL INFORMATION:
; APPLICANT: Paszty, Christopher
; APPLICANT: Gao, Yongming
; TITLE OF INVENTION: Cysteine Knot Polypeptides: Cloaked-2 Molecules and Uses Thereof
; FILE REFERENCE: 01017/37428
; CURRENT APPLICATION NUMBER: US/09/867,274
; PRIOR FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: US 60/208,550
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: US 60/223,542
; PRIOR FILING DATE: 2000-08-04
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 759
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-867-274-1

Query Match 10.8%; Score 69; DB 10; Length 759;
Best Local Similarity 100.0%; Pred. No. 6.9e-27;
Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 475 CTGGTGGCTCGTGAAGTGAAGCGCCTCACCGCTTCCACCAACGAGTCGAGCTCAAG 534
DB 517 CTGGTGGCTCGTGAAGTGAAGCGCCTCACCGCTTCCACCAACGAGTCGAGCTCAAG 576
QY 535 GACTTCGGG 543
DB 577 GACTTCGGG 585

RESULT 5
US-09-867-274-20/c
; Sequence 20, Application US/09867274
; Patent No. US20020106650A1
; GENERAL INFORMATION:
; APPLICANT: Paszty, Christopher
; APPLICANT: Gao, Yongming
; TITLE OF INVENTION: Cysteine Knot Polypeptides: Cloaked-2 Molecules and Uses Thereof
; FILE REFERENCE: 01017/37428
; CURRENT APPLICATION NUMBER: US/09/867,274
; PRIOR FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: US 60/208,550
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: US 60/223,542
; PRIOR FILING DATE: 2000-08-04
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 20
; LENGTH: 41
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial: PCR primer
US-09-867-274-20

Query Match 4.1%; Score 26; DB 10; Length 41;
Best Local Similarity 100.0%; Pred. No. 0.00059;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 611 AGCGGAGCTGGAGACGCTACTAG 636
DB 41 AGCGGAGCTGGAGACGCTACTAG 16

RESULT 6
US-09-867-274-14/c
; Sequence 14, Application US/09867274
; Patent No. US20020106650A1
; GENERAL INFORMATION:
; APPLICANT: Paszty, Christopher
; APPLICANT: Gao, Yongming
; TITLE OF INVENTION: Cysteine Knot Polypeptides: Cloaked-2 Molecules and Uses Thereof
; FILE REFERENCE: 01017/37428
; CURRENT APPLICATION NUMBER: US/09/867,274
; CURRENT FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: US 60/208,550
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: US 60/223,542
; PRIOR FILING DATE: 2000-08-04
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 14
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial: PCR primer
US-09-867-274-14

Query Match 3.9%; Score 25; DB 10; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.0021;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 198 TCCCTATGACGCCAAGATGTGTC 222
DB 25 TCCCTATGACGCCAAGATGTGTC 1

RESULT 7
US-09-867-274-19
; Sequence 19, Application US/09867274

; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepfer, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 20
; LENGTH: 19
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Primer for PCR
US-09-449-218D-20

Query Match 3.0%; Score 19; DB 4; Length 19;
Best Local Similarity 100.0%; Pred. No. 6.2;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 303 GGTGTGCTCGGCCAGTGC 321
Db 19 GGTGTGCTCGGCCAGTGC 1

RESULT 13
US-09-449-218D-25/c
; Sequence 25, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepfer, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 25
; LENGTH: 19
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Primer for PCR
US-09-449-218D-25

Query Match 3.0%; Score 19; DB 4; Length 19;
Best Local Similarity 100.0%; Pred. No. 6.2;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 303 GGTGTGCTCGGCCAGTGC 321
Db 19 GGTGTGCTCGGCCAGTGC 1

RESULT 14
US-09-449-218D-29/c
; Sequence 29, Application US/09449218D
; Patent No. 6395511

; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepfer, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 29
; LENGTH: 29
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Primer for PCR
US-09-449-218D-29

Query Match 2.8%; Score 18; DB 4; Length 29;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 619 CTGGAGAACGCCTACTAG 636
Db 29 CTGGAGAACGCCTACTAG 12

RESULT 15
US-09-449-218D-38
; Sequence 38, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepfer, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 38
; LENGTH: 31
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Primer for PCR
US-09-449-218D-38

Query Match 2.8%; Score 18; DB 4; Length 31;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 248 ACACCCGCTTCTGACAG 265
Db 14 ACACCCGCTTCTGACAG 31

Search completed: March 29, 2003, 10:47:09
Job time : 55.2086 secs

; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepker, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; TITLE OF INVENTION: BONE MINERALIZATION
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 2301
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-449-218D-7

Query Match 10.8%; Score 69; DB 4; Length 2301;
Best Local Similarity 100.0%; Pred. No. 7e-24;
Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 475 CTGGTGGCTCTGCAAGTCAAGCGCTCACCGCTTCCACAACCAAGTCGGAGCTCAAG 534
Db 528 CTGGTGGCTCTGCAAGTCAAGCGCTCACCGCTTCCACAACCAAGTCGGAGCTCAAG 587

QY 535 GACTTCGGG 543
Db 588 GACTTCGGG 596

RESULT 9
US-09-449-218D-18
; Sequence 18, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepker, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; TITLE OF INVENTION: BONE MINERALIZATION
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 18
; LENGTH: 9301
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-449-218D-18

Query Match 10.8%; Score 69; DB 4; Length 9301;
Best Local Similarity 100.0%; Pred. No. 6.4e-24;
Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 475 CTGGTGGCTCTGCAAGTCAAGCGCTCACCGCTTCCACAACCAAGTCGGAGCTCAAG 534
Db 3447 CTGGTGGCTCTGCAAGTCAAGCGCTCACCGCTTCCACAACCAAGTCGGAGCTCAAG 3506

QY 535 GACTTCGGG 543
Db 3507 GACTTCGGG 3515

RESULT 10
US-09-449-218D-9

; Sequence 9, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepker, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; TITLE OF INVENTION: BONE MINERALIZATION
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9
; LENGTH: 642
; TYPE: DNA
; ORGANISM: Cercopithecus pygerythrus
US-09-449-218D-9

Query Match 10.7%; Score 68; DB 4; Length 642;
Best Local Similarity 100.0%; Pred. No. 2.3e-23;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 475 CTGGTGGCTCTGCAAGTCAAGCGCTCACCGCTTCCACAACCAAGTCGGAGCTCAAG 534
Db 481 CTGGTGGCTCTGCAAGTCAAGCGCTCACCGCTTCCACAACCAAGTCGGAGCTCAAG 540

QY 535 GACTTCGGG 542
Db 541 GACTTCGGG 548

RESULT 11
US-09-449-218D-15
; Sequence 15, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepker, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; TITLE OF INVENTION: BONE MINERALIZATION
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 15
; LENGTH: 532
; TYPE: DNA
; ORGANISM: Bos torus
US-09-449-218D-15

Query Match 5.0%; Score 32; DB 4; Length 532;
Best Local Similarity 100.0%; Pred. No. 3.2e-06;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 475 CTGGTGGCTCTGCAAGTCAAGCGCTCACCGCTTCCACAACCAAGTCGGAGCTCAAG 506
Db 390 CTGGTGGCTCTGCAAGTCAAGCGCTCACCGCTTCCACAACCAAGTCGGAGCTCAAG 421

RESULT 12
US-09-449-218D-20/c
; Sequence 20, Application US/09449218D

```
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 45
; LENGTH: 642
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: CDS
; LOCATION: (1)..(639)
US-09-449-218D-45

Query Match          10.8%; Score 69; DB 4; Length 642;
Best Local Similarity 100.0%; Pred. No. 7e-24;
Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 475 CTGGTGGCCTCGTGAAGTGAAGCGCTCACCGCTTCCACAAACAGTCGGAGCTCAAG 534
    |||||||
Db 481 CTGGTGGCCTCGTGAAGTGAAGCGCTCACCGCTTCCACAAACAGTCGGAGCTCAAG 540
    |||||||

QY 535 GACTTCGGG 543
    |||||||
Db 541 GACTTCGGG 549

RESULT 5
US-09-449-218D-1
; Sequence 1, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepker, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 2301
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-449-218D-1

Query Match          10.8%; Score 69; DB 4; Length 2301;
Best Local Similarity 100.0%; Pred. No. 7e-24;
Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 475 CTGGTGGCCTCGTGAAGTGAAGCGCTCACCGCTTCCACAAACAGTCGGAGCTCAAG 534
    |||||||
Db 528 CTGGTGGCCTCGTGAAGTGAAGCGCTCACCGCTTCCACAAACAGTCGGAGCTCAAG 587
    |||||||

QY 535 GACTTCGGG 543
    |||||||
Db 541 GACTTCGGG 549

RESULT 6
US-09-449-218D-3
; Sequence 3, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepker, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 2301
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-449-218D-3

Query Match          10.8%; Score 69; DB 4; Length 2301;
Best Local Similarity 100.0%; Pred. No. 7e-24;
Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 475 CTGGTGGCCTCGTGAAGTGAAGCGCTCACCGCTTCCACAAACAGTCGGAGCTCAAG 534
    |||||||
Db 528 CTGGTGGCCTCGTGAAGTGAAGCGCTCACCGCTTCCACAAACAGTCGGAGCTCAAG 587
    |||||||

QY 535 GACTTCGGG 543
    |||||||
Db 541 GACTTCGGG 549

RESULT 7
US-09-449-218D-5
; Sequence 5, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepker, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 2301
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-449-218D-5

Query Match          10.8%; Score 69; DB 4; Length 2301;
Best Local Similarity 100.0%; Pred. No. 7e-24;
Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 475 CTGGTGGCCTCGTGAAGTGAAGCGCTCACCGCTTCCACAAACAGTCGGAGCTCAAG 534
    |||||||
Db 528 CTGGTGGCCTCGTGAAGTGAAGCGCTCACCGCTTCCACAAACAGTCGGAGCTCAAG 587
    |||||||

QY 535 GACTTCGGG 543
    |||||||
Db 541 GACTTCGGG 549

RESULT 8
US-09-449-218D-7
; Sequence 7, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepker, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 2301
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-449-218D-7

Query Match          10.8%; Score 69; DB 4; Length 2301;
Best Local Similarity 100.0%; Pred. No. 7e-24;
Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 475 CTGGTGGCCTCGTGAAGTGAAGCGCTCACCGCTTCCACAAACAGTCGGAGCTCAAG 534
    |||||||
Db 528 CTGGTGGCCTCGTGAAGTGAAGCGCTCACCGCTTCCACAAACAGTCGGAGCTCAAG 587
    |||||||

QY 535 GACTTCGGG 543
    |||||||
Db 541 GACTTCGGG 549
```


Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 105 GGTCATCCCGGGCTTGGAGAG 126
| | | | | | | | | | | | | | | | | |
Db 40 GGTCATCCCGGGCTTGGAGAG 61

Search completed: March 29, 2003, 10:45:09
Job time : 1374.57 secs

RPCI-24 Mouse BAC Library produced by Pieter de Jong. The library was cloned in the pTARBAC1 cloning vector at the BamHI sites using MboI partially digested male C57BL/6J DNA."

BASE COUNT 30 a 44 c 45 g 28 t
ORIGIN

Query Match 7.1%; Score 45; DB 17; Length 147;
Best Local Similarity 100.0%; Pred. No. 4e-11;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 222 CGAGTACAGCTCCGGAGCTGACCTACACCGCTTCCTGACAGA 266
|||||
DB 101 CGAGTACAGCTCCGGAGCTGACCTACACCGCTTCCTGACAGA 145
|||||

RESULT 14
BF523030 254 bp mRNA linear EST 11-DEC-2000
LOCUS
DEFINITION UI-R-C2p-rk-a-03-0-UI-r1 UI-R-C2p Rattus norvegicus cDNA clone
UI-R-C2p-rk-a-03-0-UI 5', mRNA sequence.

ACCESSION BF523030
VERSION BF523030.1 GI:11631045

KEYWORDS EST.
SOURCE Norway rat.

ORGANISM

Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

1 (bases 1 to 254)

REFERENCE
AUTHORS

Bonaldo, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery

Genome Res. 6 (9), 791-806 (1996)

97044477

CONTACT: Soares, MB

Program for Rat Gene Discovery and Mapping

University of Iowa

451 Eckstein Medical Research Building Iowa City, IA 52242, USA

Tel: 319 335 8250

Fax: 319 335 9565

Email: msoares@blue.weeg.uiowa.edu

cDNA Library Preparation: M.B. Soares Lab Clone distribution:

clones will be available through Research Genetics (www.resgen.com)
This clone is also available through the I.M.A.G.E. Consortium at
LLNL (info@image.llnl.gov). IMAGE ID= 1793134

Seq primer: M13 Forward.

Location/Qualifiers

1..254

/organism="Rattus norvegicus"

/strain="Sprague-Dawley"

/db_xref="taxon:10116"

/clone="UI-R-C2p-rk-a-03-0-UI"

/clone.lib="UI-R-C2p"

/dev_stage="adult"

/lab_host="DH10B (Life Technologies)"

/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-C2p
library is a subtracted library derived from the UI-R-C1
library, which is a subtracted library derived from the
UI-R-C0 library. The UI-R-C0 library consisted of a
mixture of individually tagged normalized libraries
constructed from rat placenta, adult lung, brain, liver,
kidney, heart, spleen, ovary, muscle, 8, 12 and 18-day
embryo. The tag is a string of 3-5 nucleotides present
between the Not I site and the oligo-dr track which allows
identification of the library of origin of a clone within
the mixture. The subtracted library (UI-R-C2p) was
constructed as follows: PCR amplified cDNA inserts from
UI-R-C1 clones from which 3' ESTs had been derived was
used as a driver in a hybridization with the UI-R-C1
library in the form of single-stranded circles. The
remaining single-stranded circles (subtracted library) was

purified by hydroxyapatite column chromatography,
converted to double-stranded circles and electroporated
into DH10B bacteria (Life Technologies) to generate the
UI-R-C2p library. This procedure has been previously
described (Bonaldo, Lennon and Soares, Genome Research 6:
791-806, 1996).

BASE COUNT 61 a 50 c 53 g 50 t
ORIGIN

Query Match 3.6%; Score 23; DB 12; Length 254;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 146 AGAACACCAACGACCATGACCGG 168
|||||

DB 186 AGAACACCAACGACCATGACCGG 208
|||||

RESULT 15

BE835032

LOCUS

DEFINITION

RC6-FN0019-290600-011-A02 FN0019 Homo sapiens cDNA, mRNA sequence.

ACCESSION BE835032

VERSION BE835032.1 GI:10267410

KEYWORDS EST.

SOURCE human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 385)

REFERENCE

AUTHORS

Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,

Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,

Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,

Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare

, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and

Simpson, A.J.

Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

20202663

COMMENT

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome

Project. This entry can be seen in the following URL

(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=RC6-FN0019-290

600-011-A02&tl=2000-06-29&tl=1)

Seq primer: puc 18 forward

High quality sequence start: 14

High quality sequence stop: 361.

Location/Qualifiers

1..385

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone.lib="FN0019"

/dev_stage="Adult"

/note="Organ: prostate normal; Vector: puc18; Site_1: SmaI

; Site_2: SmaI; A mini-library was made by cloning

products derived from ORESTES PCR (U.S. Letters Patent

application No. 196,716 - Ludwig Institute for Cancer

Research) profiles into the pUC 18 vector. Reverse

transcription of tissue mRNA and cDNA amplification were

performed under low stringency conditions."

92 a 89 c 119 g 85 t

BASE COUNT

ORIGIN

Query Match 3.5%; Score 22; DB 12; Length 385;

Best Local Similarity 100.0%; Pred. No. 6.9;

Soares Lab Clone distribution: clones will be available through Research Genetics (www.resgen.com) The following repetitive elements were found in this cdna sequence: 22-42,
>GC rich#Low complexity
Seq primer: M13 Forward
POLYA=No.

FEATURES

source

Location/Qualifiers

1. .472
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-BJ2-bpm-b-02-0-UI"
/clone_lib="UI-R-BJ2"
/lab_host="DH10B (Life Technologies)"
/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-BJ2 library is a subtracted library derived from the following tissues: heart, atrium at 15 dpc, ventricle at 16.5 dpc, atrium at 16.5 dpc, ventricle at 13 dpc, ventricle at 15 dpc, AV canal at 15 dpc. For a detailed description of the library from which this clone was derived, please visit our web site at ratest.eng.uiowa.edu. The subtraction has been previously described in (Bonaldo, Lennon and Soares, Genome Research 6:791-806, 1996)
TAG_SEQ=None found"

BASE COUNT 103 a 148 c 147 g 74 t

ORIGIN

Query Match 7.9%; Score 50; DB 12; Length 472;

Best Local Similarity 100.0%; Pred. No. 1.7e-13;

Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 587 GCGCCGGGAGCCAAAGCCAGCGGAGCTGGAGAACGCTACTAG 636

Db 33 GCGCCGGGAGCCAAAGCCAGCGGAGCTGGAGAACGCTACTAG 82

RESULT 12

BF407514

LOCUS

DEFINITION UI-R-BJ2-bqt-b-08-0-UI.s1 UI-R-BJ2 Rattus norvegicus cDNA clone

UI-R-BJ2-bqt-b-08-0-UI 3', mRNA sequence.

ACCESSION BF407514

VERSION BF407514.1 GI:11395489

KEYWORDS EST.

SOURCE Norway rat.

ORGANISM

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

1 (bases 1 to 532)

Normalization and subtraction: two approaches to facilitate gene

discovery

Genome Res. 6 (9), 791-806 (1996)

97044477

Contact: Soares, MB

Program for Rat Gene Discovery and Mapping

University of Iowa

451 Eckstein Medical Research Building Iowa City, IA 52242, USA

Tel: 319 335 8250

Fax: 319 335 9565

Email: msoares@blue.weeg.uiowa.edu

Oligo-dt track not found, Not I site shown in beginning of sequence

is likely internal to the message. cDNA Library Preparation: M.B.

Soares Lab Clone distribution: clones will be available through

Research Genetics (www.resgen.com) The following repetitive

elements were found in this cdna sequence: 22-42,

>GC rich#Low complexity

Seq primer: M13 Forward

POLYA=No.

FEATURES

source

Location/Qualifiers

1. .532

/organism="Rattus norvegicus"

/strain="Sprague-Dawley"

/db_xref="taxon:10116"

/clone="UI-R-BJ2-bqt-b-08-0-UI"

/clone_lib="UI-R-BJ2"

/lab_host="DH10B (Life Technologies)"

/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-BJ2 library is a subtracted library derived from the following tissues: heart, atrium at 15 dpc, ventricle at 16.5 dpc, atrium at 16.5 dpc, ventricle at 13 dpc, ventricle at 15 dpc, AV canal at 15 dpc. For a detailed description of the library from which this clone was derived, please visit our web site at ratest.eng.uiowa.edu. The subtraction has been previously described in (Bonaldo, Lennon and Soares, Genome Research 6:791-806, 1996)
TAG_SEQ=None found"

BASE COUNT 111 a 160 c 167 g 94 t

ORIGIN

Query Match 7.9%; Score 50; DB 12; Length 532;

Best Local Similarity 100.0%; Pred. No. 1.7e-13;

Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 587 GCGCCGGGAGCCAAAGCCAGCGGAGCTGGAGAACGCTACTAG 636

Db 33 GCGCCGGGAGCCAAAGCCAGCGGAGCTGGAGAACGCTACTAG 82

RESULT 13

AZ746508

LOCUS

DEFINITION RPCI-24-79P16.TJB RPCI-24 Mus musculus genomic clone RPCI-24-79P16,

DNA sequence.

ACCSSION AZ746508

VERSION AZ746508.1 GI:12529424

KEYWORDS GSS.

SOURCE house mouse.

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 147)

REFERENCE 1 (bases 1 to 147)

AUTHORS Tsegaye, G., Geer, K., Krol, M., Shvartsbeyn, A., Gebregorgis, E.,

Zhao, S., Nierman, W., Malek, J., Shatsman, S., Akinret, B., Levins, M.,

Russell, D., de Jong, P., and Fraser, C.M.

Mouse BAC End Sequences from Library RPCI-24

Unpublished (1999)

Other_GSSs: RPCI-24-79P16.TVB

Contact: Shaying Zhao

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: szhao@tigr.org

Clones are derived from the mouse BAC library RPCI-24. For BAC

library availability, please contact Pieter de Jong

(pdejong@mail.cho.org). Clones may be purchased from BACPAC

Resources (<http://www.chori.org/bacpac/orderingframe.htm>). BAC end

page: http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html

Plate: 79 row: P column: 16

Seq primer: SP6

Class: BAC ends.

Location/Qualifiers

1. .147

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="RPCI-24-79P16"

/clone_lib="RPCI-24"

/sex="Male"

/cell_type="Spleen/Brain"

/notes="Vector: pTARBAC1; Site_1: BamHI; Site_2: BamHI;

Plate: 0191 row: D column: 21
 Seq primer: CGTTGTAAACACGCGCCAGT
 Class: plasmid ends
 High quality sequence stop: 627.
 Location/Qualifiers

FEATURES

source

```

1..627
  /organism="Mus musculus"
  /strain="C57BL/6J"
  /db_xref="taxon:10090"
  /clone="UUG2M0191D21"
  /clone_lib="Mouse 10kb plasmid UUGC1M library"
  /sex="Male"
  /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
  /note="Vector: pWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
  BASE COUNT 168 a 157 c 176 g 126 t
  ORIGIN

```

```

Query Match      10.7%; Score 68; DB 17; Length 627;
Best Local Similarity 100.0%; Pred. No. 3.1e-22;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 213 AGATGTGTCGAGTACAGTGGCGGAGCTGCACCTACACCCGCTTCTGACAGCGGCC 272
      |||||
Db 560 AGATGTGTCGAGTACAGTGGCGGAGCTGCACCTACACCCGCTTCTGACAGCGGCC 619

QY 273 ATGCCGCA 280
      |||||
Db 620 ATGCCGCA 627

```

```

RESULT 10
AW434705
LOCUS      AW434705          385 bp      mRNA      linear      EST 11-FEB-2000
DEFINITION UI-R-BJ0p-afz-c-02-0-UI.s1 UI-R-BJ0p Rattus norvegicus cDNA clone
            UI-R-BJ0p-afz-c-02-0-UI 3', mRNA sequence.
ACCESSION  AW434705
VERSION     AW434705.1 GI:6966012
KEYWORDS    EST.
SOURCE      Norway rat.
            Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.

```

```

REFERENCE   1 (bases 1 to 385)
AUTHORS    Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE       Normalization and subtraction: two approaches to facilitate gene
            discovery
JOURNAL     Genome Res. 6 (9), 791-806 (1996)
MEDLINE     9704477
COMMENT     Contact: Soares, MB
            Program for Rat Gene Discovery and Mapping
            University of Iowa
            451 Eckstein Medical Research Building Iowa City, IA 52242, USA
            Tel: 319 335 8250
            Fax: 319 335 9565

```

Email: msoares@blue.weeg.uiowa.edu
 Oligo-dT track not found, Not I site shown in beginning of sequence
 is likely internal to the message. cDNA Library Preparation: M.B.
 Soares Lab Clone distribution: clones will be available through
 Research Genet.cs (www.resgen.com) The following repetitive
 elements were found in this cDNA sequence: 21-41,
 >GC-rich#Low complexity
 Seq primer: M13 Forward
 POLYA=No.

FEATURES

source

```

1..385
  /organism="Rattus norvegicus"
  /strain="Sprague-Dawley"
  /db_xref="taxon:10116"
  /clone="UI-R-BJ0p-afz-c-02-0-UI"
  /clone_lib="UI-R-BJ0p"
  /dev_stage="adult"
  /lab_host="DH10B (Life Technologies)"
  /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
  polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-BJ0p
  library is a subtracted library derived from the UI-R-AA1,
  UI-R-AB1, UI-R-AC1, UI-R-AD1, UI-R-AE1, UI-R-AF1, and
  UI-R-AG1 libraries. These libraries represent tissues from
  rat atrium at 16.5 dpc, ventricle at 16.5 dpc, AV canal
  at 16.5 dpc, atrium at 15 dpc, ventricle at 15 dpc, AV
  canal at 15 dpc, and ventricle at 13 dpc. The tag is a
  string of 5-6 nucleotides present between the Not I site
  and the oligo-dT track. The library was constructed as
  described by Bonaldo, Lennon and Soares, Genome Research
  6: 791-806, 1996.
  TAG LAB=UI-R-BJ0p
  TAG TSSUE=AV canal at 16.5 dpc
  TAG SEQ=GAACC"
  BASE COUNT 75 a 137 c 122 g 61 t
  ORIGIN

```

```

Query Match      7.9%; Score 50; DB 10; Length 385;
Best Local Similarity 100.0%; Pred. No. 1.7e-13;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 587 GCGCCCGGAGCCAAAGCCAGCGGAGCTGGAGAACGCTACTAG 636
      |||||
Db 32 GCGCCCGGAGCCAAAGCCAGCGGAGCTGGAGAACGCTACTAG 81

```

```

RESULT 11
BF420024
LOCUS      BF420024          472 bp      mRNA      linear      EST 28-NOV-2000
DEFINITION UI-R-BJ2-bpm-b-02-0-UI.s1 UI-R-BJ2 Rattus norvegicus cDNA clone
            UI-R-BJ2-bpm-b-02-0-UI 3', mRNA sequence.
ACCESSION  BF420024
VERSION     BF420024.1 GI:11408013
KEYWORDS    EST.
SOURCE      Norway rat.
            Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.

```

```

REFERENCE   1 (bases 1 to 472)
AUTHORS    Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE       Normalization and subtraction: two approaches to facilitate gene
            discovery
JOURNAL     Genome Res. 6 (9), 791-806 (1996)
MEDLINE     9704477
COMMENT     Contact: Soares, MB
            Program for Rat Gene Discovery and Mapping
            University of Iowa
            451 Eckstein Medical Research Building Iowa City, IA 52242, USA
            Tel: 319 335 8250
            Fax: 319 335 9565
            Email: msoares@blue.weeg.uiowa.edu
            Oligo-dT track not found, Not I site shown in beginning of sequence
            is likely internal to the message. cDNA Library Preparation: M.B.

```

library is a subtracted library derived from the UI-R-C1 library, which is a subtracted library derived from the UI-R-C0 library. The UI-R-C0 library consisted of a mixture of individually tagged normalized libraries constructed from rat placenta, adult lung, brain, liver, kidney, heart, spleen, ovary, muscle, 8, 12 and 18-day embryo. The tag is a string of 3-5 nucleotides present between the Not I site and the oligo-dT track which allows identification of the library of origin of a clone within the mixture. The subtracted library (UI-R-C2p) was constructed as follows: PCR amplified cDNA inserts from UI-R-C1 clones from which 3' ESTs had been derived was used as a driver in a hybridization with the UI-R-C1 library in the form of single-stranded circles. The remaining single-stranded circles (subtracted library) was purified by hydroxyapatite column chromatography, converted to double-stranded circles and electroporated into DH10B bacteria (Life Technologies) to generate the UI-R-C2p library. This procedure has been previously described (Bonaldo, Lennon and Soares, Genome Research 6: 791-806, 1996)"

BASE COUNT 54 a 120 c 131 g 54 t 1 others
ORIGIN

Query Match 15.4%; Score 98; DB 9; Length 360;
Best Local Similarity 100.0%; Pred. No. 7.4e-37;
Matches 98; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 445 GCGCGCGCGCGCTCGCGCAGGTGCTGTGGCTCTGCAAGTGCAGCGCTC 504
|||||
Db 117 GCGCGCGCGCGCTCGCGCAGGTGCTGTGGCTCTGCAAGTGCAGCGCTC 58
|||||

QY 505 ACCCGCTTCACAAACAGTCGAGCTCAAGGACTTCGG 542
|||||
Db 57 ACCCGCTTCACAAACAGTCGAGCTCAAGGACTTCGG 20
|||||

RESULT 8
AII13131/c
LOCUS AII13131 315 bp mRNA linear EST 11-FEB-1999
DEFINITION UI-R-C2p-nr-f-06-0-UI.s1 UI-R-C2p Rattus norvegicus cDNA clone
UI-R-C2p-nr-f-06-0-UI 3', mRNA sequence.

ACCESSION AII13131
VERSION AII13131.1 GI:3513080
KEYWORDS EST.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE 1 (bases 1 to 315)
Bonaldo,M.F., Lennon,G. and Soares,M.B.
Normalization and subtraction: two approaches to facilitate gene discovery
Genome Res. 6 (9), 791-806 (1996)

JOURNAL 9704477
MEDLINE
COMMENT Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: msoares@blue.weeg.uiowa.edu
Oligo-dT track not found, Not I site shown in beginning of sequence
is likely internal to the message. cDNA Library Preparation: M.
Fatima Bonaldo, Ph.D. Clone distribution: clones will be available
through Research Genetics The following repetitive elements were
found in this cDNA sequence: 97-125, >GC_rich#Low_complexity
Seq primer: M13 Forward.

FEATURES
source
1..315
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"

/db xref="taxon:10116"
/clone="UI-R-C2p-nr-f-06-0-UI"
/clone_lib="UI-R-C2p"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies)"
/notes="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-C2p library is a subtracted library derived from the UI-R-C1 library, which is a subtracted library derived from the UI-R-C0 library. The UI-R-C0 library consisted of a mixture of individually tagged normalized libraries constructed from rat placenta, adult lung, brain, liver, kidney, heart, spleen, ovary, muscle, 8, 12 and 18-day embryo. The tag is a string of 3-5 nucleotides present between the Not I site and the oligo-dT track which allows identification of the library of origin of a clone within the mixture. The subtracted library (UI-R-C2p) was constructed as follows: PCR amplified cDNA inserts from UI-R-C1 clones from which 3' ESTs had been derived was used as a driver in a hybridization with the UI-R-C1 library in the form of single-stranded circles. The remaining single-stranded circles (subtracted library) was purified by hydroxyapatite column chromatography, converted to double-stranded circles and electroporated into DH10B bacteria (Life Technologies) to generate the UI-R-C2p library. This procedure has been previously described (Bonaldo, Lennon and Soares, Genome Research 6: 791-806, 1996)"

BASE COUNT 46 a 110 c 116 g 43 t
ORIGIN

Query Match 11.2%; Score 71; DB 9; Length 315;
Best Local Similarity 100.0%; Pred. No. 9.9e-24;
Matches 71; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 313 GGCAGTGCAGCGCGCGCTCTGCCAACGCCATCGGCGCGTGAAGTGGTGGCG 372
|||||
Db 249 GGCAGTGCAGCGCGCGCTCTGCCAACGCCATCGGCGCGTGAAGTGGTGGCG 190
|||||

QY 373 CCGAACGGACC 383
|||||
Db 189 CCGAACGGACC 179
|||||

RESULT 9
AZ876400
LOCUS AZ876400 627 bp DNA linear GSS 21-FEB-2001
DEFINITION 2M0191D21F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC2M0191D21 F, DNA sequence.

ACCESSION AZ876400
VERSION AZ876400.1 GI:13087273
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 627)

REFERENCE 1 (bases 1 to 627)
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D.,Weiss,R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
84112, USA

Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00

canal at 16.5 dpc, atrium at 15 dpc, ventricle at 15 dpc, AV canal at 15 dpc, ventricle at 13 dpc, and adult heart. For a detailed description of the library from which this clone was derived, please visit our web site at ratseq.eng.uiowa.edu. The subtraction has been previously described in (Bonaldo, Lennon and Soares, Genome Research 6:791-806, 1996)
TAG_SEQ=None found"

BASE COUNT 41 a 101 c 109 g 40 t
ORIGIN

Query Match 16.4%; Score 104; DB 10; Length 291;
Best Local Similarity 99.4%; Pred. No. 8.7e-40;
Matches 154; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 388 TTCGGCTGATCCCGGATCGCTACCGCGCGAGCGGTGCGAGCTCTGTGCCCGCGGGGC 447
DB 174 TTCGGCTGATCCCGGATCGCTACCGCGCGAGCGGTGCGAGCTCTGTGCCCGCGGGGC 115
QY 448 GCGGCGCGCGCTCGCGCAAGTGGCTGTGGTGGCTCTGTGCAAGTGAAGCGCTCACC 507
DB 114 GCGGCGCGCGCTCGCGCAAGTGGCTGTGGTGGCTCTGTGCAAGTGAAGCGCTCACC 55
QY 508 CGCTTCCCAACAGTCGGAGCTCAAGGACTTCGG 542
DB 54 CGCTTCCCAACAGTCGGAGCTCAAGGACTTCGG 20

RESULT 6

BE101082/c
LOCUS BE101082 419 bp mRNA linear EST 13-JUN-2000
DEFINITION UI-R-BJ1-aty-h-10-0-UI.s1 UI-R-BJ1 Rattus norvegicus cDNA clone
UI-R-BJ1-aty-h-10-0-UI 3', mRNA sequence.

ACCESSION BE101082
VERSION BE101082.1 GI:8492989

KEYWORDS EST.

SOURCE Norway rat.

ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE 1 (bases 1 to 419)

AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.

TITLE Normalization and subtraction: two approaches to facilitate gene discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)

MEDLINE 9704477

COMMENT

Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565

Email: msoares@blue.weeg.uiowa.edu
Oligo-dT track not found. Not 1 site shown in beginning of sequence
is likely internal to the message. cDNA Library Preparation: M.B.
Soares Lab Clone distribution: Clones will be available through
Research Genetics (www.resgen.com) The following repetitive
elements were found in this cDNA sequence: 97-125,
>GC rich#Low complexity

Seq primer: M13 Forward

POLYA=No.

Location/Qualifiers

FEATURES

source

1..419

/organism="Rattus norvegicus"

/strain="Sprague-Dawley"

/db_xref="taxon:10116"

/clone="UI-R-BJ1-aty-h-10-0-UI"

/clone_lib="UI-R-BJ1"

/lab_host="DH10B (Life Technologies)"

/note="Vector: pRT3D-Pac (Pharmacia) with a modified

polylinker; Site 1: Not 1; Site 2: Eco RI; The UI-R-BJ1

library is a subtracted library derived from the following

tissues: atrium at 16.5 dpc, ventricle at 16.5 dpc, AV canal at 16.5 dpc, atrium at 15 dpc, ventricle at 15 dpc, AV canal at 15 dpc, ventricle at 13 dpc, and adult heart. For a detailed description of the library from which this clone was derived, please visit our web site at ratseq.eng.uiowa.edu. The subtraction has been previously described in (Bonaldo, Lennon and Soares, Genome Research 6:791-806, 1996)
TAG_SEQ=None found"

BASE COUNT 56 a 135 c 155 g 73 t
ORIGIN

Query Match 16.4%; Score 104; DB 10; Length 419;
Best Local Similarity 99.4%; Pred. No. 9.1e-40;
Matches 154; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 388 TTCGGCTGATCCCGGATCGCTACCGCGCGAGCGGTGCGAGCTCTGTGCCCGCGGGGC 447
DB 174 TTCGGCTGATCCCGGATCGCTACCGCGCGAGCGGTGCGAGCTCTGTGCCCGCGGGGC 115
QY 448 GCGGCGCGCGCTCGCGCAAGTGGCTGTGGTGGCTCTGTGCAAGTGAAGCGCTCACC 507
DB 114 GCGGCGCGCGCTCGCGCAAGTGGCTGTGGTGGCTCTGTGCAAGTGAAGCGCTCACC 55
QY 508 CGCTTCCCAACAGTCGGAGCTCAAGGACTTCGG 542
DB 54 CGCTTCCCAACAGTCGGAGCTCAAGGACTTCGG 20

RESULT 7

AI556282/c

LOCUS AI556282 360 bp mRNA linear EST 23-MAR-1999

DEFINITION UI-R-C2p-rk-a-03-0-UI.s1 UI-R-C2p Rattus norvegicus cDNA clone
UI-R-C2p-rk-a-03-0-UI 3', mRNA sequence.

ACCESSION AI556282

VERSION AI556282.1 GI:4488645

KEYWORDS EST.

SOURCE Norway rat.

ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

REFERENCE 1 (bases 1 to 360)

AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.

TITLE Normalization and subtraction: two approaches to facilitate gene discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)

MEDLINE 9704477

COMMENT

Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565

Email: msoares@blue.weeg.uiowa.edu
Oligo-dT track not found. Not 1 site shown in beginning of sequence
is likely internal to the message. cDNA Library Preparation: M.
Patina Bonaldo, Ph.D. Clone distribution: Clones will be available
through Research Genetics (www.resgen.com) The following repetitive
elements were found in this cDNA sequence: 97-125,
>GC rich#Low complexity

Seq primer: M13 Forward

Location/Qualifiers

FEATURES

source

1..360

/organism="Rattus norvegicus"

/strain="Sprague-Dawley"

/db_xref="taxon:10116"

/clone="UI-R-C2p-rk-a-03-0-UI"

/clone_lib="UI-R-C2p"

/dev_stage="adult"

/lab_host="DH10B (Life Technologies)"

/note="Vector: pRT3D-Pac (Pharmacia) with a modified

polylinker; Site 1: Not 1; Site 2: Eco RI; The UI-R-C2p

Fax: 81-45-503-9216

Email: genome-res@gs.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh
M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
Waghi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,
Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura
S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and
Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara
Y. and Hayashizaki, Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Kondo, S., Shinagawa, A., Saito, T., Kiyosawa, H., Yamanaka, I., Aizawa
K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and
Hayashizaki, Y.
Computational Analysis of Full-Length Mouse cDNAs Compared with
Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.
e mouse tissues.

FEATURES

Location/Qualifiers
1. .667
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="A53000IM15"
/clone_lib="RIKEN full-length enriched, adult male aorta
and vein"
/sex="male"
/tissue_type="aorta and vein"
/dev_stage="adult"
/lab_host="DH10B"
/note="Site 1: Sall; Site 2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'
GAGAGAGAGAGATCCAGAGCTCTTTTCTTTTCTTNN 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. cDNA went through one round of normalization
to Rot = 20.0 and subtraction to Rot = 459.0. Second
strand cDNA was prepared with the primer adapter of
sequence [5' GAGAGAGATTCGAGTAAATTAATATCCCCCCCCCC
3']. cDNA was cleaved with XhoI and BamHI. Vector: a
modified pBluescript KS(+) after bulk excision from Lambda
FLC I."

BASE COUNT 119 a 234 c 212 g 100 t 2 others
ORIGIN

Query Match 55.0%; Score 350; DB 10; Length 667;
Best Local Similarity 99.4%; Pred. No. 2.2e-159;
Matches 500; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGCAGCCCTCACTAGCCCGTGCCTCATCTGCTACTTGTGCGAGCTGCCCTCTGTGCT 60
DB 41 ATGCAGCCCTCACTAGCCCGTGCCTCATCTGCTACTTGTGCGAGCTGCCCTCTGTGCT 100
QY 61 GTGAGGGCCAGGGGTGGCAAGCTTTCAGAAATGATGCGACAGAGTATCCACGGGCTT 120
DB 101 GTGAGGGCCAGGGGTGGCAAGCTTTCAGAAATGATGCGACAAAGGTATCCACGGGCTT 160
QY 121 GGAGATACCCGAGCTCTCTCTGAGACACACAGACCATGAACCGGGCGAGATGGA 180

Db 161 GGAGAGTACCCCGAGCCTCTCTTGAGAAACAACAGACCATGAACCGGGCGGAGAAATGGA 220
QY 181 GCGACACCTCCCAACCATCCCTATGAGCCCAAGATGTGTCGAGTACAGTCCGCGGAG 240
Db 221 GCGACACCTCCCAACCATCCCTATGAGCCCAAGATGTGTCGAGTACAGTCCGCGGAG 280
QY 241 CTGCACTACACCCGCTTCTCTGACAGAGCGGCCATGCGCGAGCGCAAGCCGGTCAACGAG 300
Db 281 CTGCACTACACCCGCTTCTCTGACAGAGCGGCCATGCGCGAGCGCAAGCCGGTCAACGAG 340
QY 301 TTGTTGTCTCCGCGCAGTGGCGCGCGCGCGCTGCTGCCACAGCGCATCGGGCGGTG 360
Db 341 TTGTTGTCTCCGCGCAGTGGCGCGCGCGCGCTGCTGCCACAGCGCATCGGGCGGTG 400
QY 361 AAGTGGTGGCGCGCGAGCGGACCGGATTTCCGCTGTCATCCCGATCGCTACCGCGCGGAG 420
Db 401 AAGTGGTGGCGCGCGAGCGGACCGGATTTCCGCTGTCATCCCGATCGCTACCGCGCGGAG 460
QY 421 CGGTTGCAAGTGTGTTGCCCGCGCGCGCGCGCTGCGCAAGGTGCTGCTGCTG 480
Db 461 CGGTTGCAAGTGTGTTGCCCGCGCGCGCGCGCTGCGCAAGGTGCTGCTGCTG 520
QY 481 GCCTCGTGCAGTGCAGAGCGCT 503
Db 521 GCCTCGTGCAGTGCAGAGCGCT 543

RESULT 5
BEL11224/c
LOCUS BEL11224 291 bp mRNA linear EST 13-JUN-2000
DEFINITION UI-R-BJ1-auw-b-04-0-UI.s1 UI-R-BJ1 Rattus norvegicus cDNA clone
UI-R-BJ1-auw-b-04-0-UI 3', mRNA sequence.
ACCESSION BEL11224
VERSION BEL11224.1 GI:8503329
KEYWORDS EST.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE 1 (bases 1 to 291)
AUTHORS Ronaldo, M.F., Lennon, G. and Soares, M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
COMMENT Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: msoares@blue.weeg.uiowa.edu

Oligo-dt track not found, Not 1 site shown in beginning of sequence
is likely internal to the message. cDNA library preparation: M.B.
Soares Lab Clone distribution: clones will be available through
Research Genetics (www.resgen.com) The following repetitive
elements were found in this cDNA sequence: 97-125,
>C-rich#Low complexity
Seq primer: M13 Forward
POLYA=No.

FEATURES
source
1. .291
Location/Qualifiers
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-BJ1-auw-b-04-0-UI"
/clone_lib="UI-R-BJ1"
/lab_host="DH10B (Life Technologies)"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not 1; Site 2: Eco RI; The UI-R-BJ1
library is a subtracted library derived from the following
tissues: atrium at 16.5 dpc, ventricle at 16.5 dpc, AV

FEATURES

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 29, 2003, 01:29:00 ; Search time 1369.57 Seconds
(without alignments)
7520.874 Million cell updates/sec

Title: US-09-867-274-3

Perfect score: 636

Sequence: 1 atgcagccctactagcccc.....agctggagacgcctactag 636

Scoring table: OLIGO_NUC

Gapop_60.0 , Gapext 60.0

Searched: 16154066 seqs, 8097743376 residues

Word size : 0

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

EST:*

1:	em_estba:*
2:	em_esthum:*
3:	em_estin:*
4:	em_estmu:*
5:	em_estov:*
6:	em_estpl:*
7:	em_estro:*
8:	em_htc:*
9:	gb_est1:*
10:	gb_est2:*
11:	gb_htc:*
12:	gb_est3:*
13:	gb_est4:*
14:	gb_est5:*
15:	em_estfun:*
16:	em_estom:*
17:	gb_gss:*
18:	em_gss_hum:*
19:	em_gss_inv:*
20:	em_gss_pln:*
21:	em_gss_vrt:*
22:	em_gss_fun:*
23:	em_gss_nam:*
24:	em_gss_mus:*
25:	em_gss_other:*
26:	em_gss_pro:*
27:	em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	636	100.0	1990	11 AK017295	AK017295 Mus muscu
2	583	91.7	623	10 BB637315	BB637315 BB637315
3	469	73.7	511	10 BB638050	BB638050 BB638050
4	350	55.0	667	10 BB636457	BB636457 BB636457
C 5	104	16.4	291	10 BB11224	BB11224 UI-R-BJ1-
C 6	104	16.4	419	10 BB101082	BB101082 UI-R-BJ1-

C 7	98	15.4	360	9 AI556282	AI556282 UI-R-C2p-
C 8	71	11.2	315	9 AI113131	AI113131 UI-R-C2p-
C 9	68	10.7	627	17 AZ876400	AZ876400 2M019D21
10	50	7.9	385	10 AW434705	AW434705 UI-R-BJ0P
11	50	7.9	472	12 BF420024	BF420024 UI-R-BJ2-
12	50	7.9	532	12 BF407514	BF407514 UI-R-BJ2-
13	45	7.1	147	17 AZ746508	AZ746508 RPK1-24-7
14	23	3.6	254	12 BF523030	BF523030 UI-R-C2p-
15	22	3.5	385	12 BE835032	BE835032 RCG-FN001
16	22	3.5	355	17 AQ171546	AQ171546 HS 3088 B
C 17	21	3.3	301	10 BB229435	BB229435 BB229435
C 18	21	3.3	315	9 AI356622	AI356622 qy16a11.x
C 19	21	3.3	351	10 BB115441	BB115441 BB115441
C 20	21	3.3	422	13 BM668009	BM668009 UI-E-DW0-
C 21	21	3.3	458	9 AI096436	AI096436 qb59d11.x
C 22	21	3.3	460	14 BM689096	BM689096 UI-E-CQ1-
C 23	21	3.3	477	10 AW293881	AW293881 UI-H-BW0-
24	21	3.3	497	17 AQ266759	AQ266759 RPK111-76
25	21	3.3	552	14 BM696978	BM696978 UI-E-DW0-
26	21	3.3	646	17 CNS05C5X	AL330702 Tetradon
27	21	3.3	787	12 BG820544	BG820544 602782887
28	20	3.1	248	12 BG050913	BG050913 FMI 72 GO
C 29	20	3.1	249	17 AZ576934	AZ576934 03c06_Sro
30	20	3.1	294	10 AW335927	AW335927 21812 MAR
31	20	3.1	296	10 AW353570	AW353570 33856 MAR
32	20	3.1	298	10 AW335931	AW335931 21806 MAR
33	20	3.1	426	13 BI537845	BI537845 428108 MA
C 34	20	3.1	434	9 AA974943	AA974943 0M02a01.B
35	20	3.1	537	10 AW461827	AW461827 BF230006B
36	20	3.1	622	12 BG851121	BG851121 1024030D0
C 37	20	3.1	935	13 BM451086	BM451086 AGENCOURT
C 38	19	3.0	243	14 BQ469230	BQ469230 HM03M14r
C 39	19	3.0	259	12 BG604260	BG604260 WHE2505.B
C 40	19	3.0	282	10 BB386370	BB386370 BB386370
C 41	19	3.0	308	10 BB205564	BB205564 BB205564
C 42	19	3.0	344	10 BB626832	BB626832 uul1h10.y
C 43	19	3.0	328	10 BB536166	BB536166 BB536166
C 44	19	3.0	424	12 BF605457	BF605457 271850 MA
C 45	19	3.0	432	14 R68738	R68738 Yil4g04.r1

ALIGNMENTS

RESULT 1
AK017295
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE

AK017295
Mus musculus 6 days neonate head cDNA, RIKEN full-length enriched library, clone:5430411E23:sclerostin, full insert sequence.
AK017295
GI:12856464
HTC; CAP trapper.
Mus musculus (strain:C57BL/6J) 6 days neonate head cDNA to mRNA, clone_lib:RIKEN full-length enriched mouse cDNA library
Clone:5430411E23.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
10349636
2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159
3

RESULT 15

AAA29061
ID AAA29061 standard; cDNA; 2301 BP.

XX AC AAA29061;

XX DT 12-SEP-2000 (first entry)

XX DE Mutant human TGF-beta binding protein (BEER) cDNA.

XX KW osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
KW BEER; chromosome 17q12-21; gene therapy; antisense therapy; fracture;
XX bone mineralization; mutant; sclerosteosis; ss.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers
FT CDS 48..119
FT /tag= a
FT /note= "mutant BEER created by sclerosteosis
FT nonsense mutation"

XX PN WO200032773-A1.

XX PD 08-JUN-2000.

XX PF 24-NOV-1999; 99WO-US27990.

XX PR 27-NOV-1998; 98US-0110283.

XX PA (DARW-) DARWIN DISCOVERY LTD.

XX PI Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepker BW;
PI Van Ness J, Winkler DG;

XX DR WPI; 2000-41321/35.

XX P-PSDB; AAY96435.

XX PT Nucleic acids (I) encoding a transforming growth factor beta binding
PT protein, useful for identifying agents for treating osteopenia,
PT osteoporosis and fractures

XX PS Claim 27; Page 117-118; 162pp; English.

XX CC This cDNA encodes a mutant human transforming growth factor-beta
CC (TGF-beta) binding protein (BEER) produced as a result of a nonsense
CC mutation in the BEER coding sequence (C to T mutation at position 117).
CC This mutation has been linked to sclerosteosis in humans by linkage
CC analysis of an Afrikaaner population in South Africa. The hBEER gene has
CC been localized to the chromosome 17q12-21. The cDNA and protein may be
CC used for prevention, treatment and diagnosis of diseases associated with
CC inappropriate BEER expression. For example, they may be used to treat
CC disorders associated with decreased TGF-beta BP expression. The cDNA or
CC vectors may be administered to treat diseases by rectifying mutations or
CC deletions in a patient's genome that affect the activity of BEER by
CC expressing inactive proteins or to supplement the patients own production
CC of BEER polypeptides. The nucleic acids may be used for recombinant
CC production of BEER, gene therapy, antisense therapy, as probes for
CC diagnostic assays and for functional studies. BEER may be used to raise
CC antibodies and for identification of BEER modulators. BEER antagonists
CC may be used to increase bone mineral content for the treatment of
CC disorders such as osteopenia, osteoporosis, fractures and other
CC disorders associated with low mineral content.

XX SQ Sequence 2301 BP; 568 A; 634 C; 614 G; 485 T; 0 other;

Query Match 10.8%; Score 69; DB 21; Length 2301;

Best Local Similarity 100.0%; Pred. No. 4.2e-23;

Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 475 CTGGTGGCCTCGTGAAGTGCAAGCGCCTCACCCGCTTCCACAACCAAGTCGGAGCTCAAG 534

Db 528 CTGGTGGCCTCGTGAAGTGCAAGCGCCTCACCCGCTTCCACAACCAAGTCGGAGCTCAAG 587

QY 535 GACTTCGGG 543

Db 588 GACTTCGGG 596

Search completed: March 29, 2003, 08:39:31
Job time : 213.734 secs

```
XX 12-SEP-2000 (first entry)
XX Human TGF-beta binding protein (BEER) cDNA.
XX
XX osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
XX BEER; chromosome 17q12-21; gene therapy; antisense therapy; fracture;
XX bone mineralization; ss.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 48..689
XX FT /*tag= a
XX FT /product= TGF-beta_binding_protein
XX
XX W0200032773-A1.
XX
XX 08-JUN-2000.
XX
XX 24-NOV-1999; 99WO-US27990.
XX
XX 27-NOV-1998; 98US-0110283.
XX (DARW-) DARWIN DISCOVERY LTD.
XX
XX Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepers BW;
XX Van Ness J, Winkler DG;
XX WPI; 2000-412321/35.
XX P-PSDB; AAY96429.
XX
XX Nucleic acids (1) encoding a transforming growth factor beta binding
XX protein, useful for identifying agents for treating osteopenia,
XX osteoporosis and fractures
XX
XX Claim 1; Page 114-115; 162pp; English.
XX
XX This cDNA encodes a human transforming growth factor-beta (TGF-beta)
XX binding protein designated BEER. The hBEER gene has been localized
XX to the chromosome 17q12-21. The cDNA and protein may be used for
XX prevention, treatment and diagnosis of diseases associated with
XX inappropriate BEER expression. For example, they may be used to treat
XX disorders associated with decreased TGF-beta BP expression. The cDNA or
XX vectors may be administered to treat diseases by rectifying mutations or
XX deletions in a patient's genome that affect the activity of BEER by
XX expressing inactive proteins or to supplement the patient's own production
XX of BEER polypeptides. The nucleic acids may be used for recombinant
XX production of BEER, gene therapy, antisense therapy, as probes for
XX diagnostic assays and for functional studies. BEER may be used to raise
XX antibodies and for identification of BEER modulators. BEER antagonists
XX may be used to increase bone mineral content for the treatment of
XX disorders such as osteopenia, osteoporosis, fractures and other
XX disorders associated with low mineral content.
XX
XX Sequence 2301 BP; 568 A; 635 C; 614 G; 484 T; 0 other;
XX
XX Query Match 10.8%; Score 69; DB 21; Length 2301;
XX Best Local Similarity 100.0%; Pred. No. 4.2e-23;
XX Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 475 CTGGTGGCTCGTGAAGTGCAGCGCCTCACCGCTTCCACACACGATCGGAGCTCAAG 534
Db 528 CTGGTGGCTCGTGAAGTGCAGCGCCTCACCGCTTCCACACACGATCGGAGCTCAAG 587
QY 535 GACTTCGGG 543
Db 588 GACTTCGGG 596
RESULT 14
AAA29056
ID AAA29056 standard; cDNA; 2301 BP.
```

```
XX AAA29056;
XX
XX 12-SEP-2000 (first entry)
XX
XX Human TGF-beta binding protein (BEER) variant V10I cDNA.
XX
XX osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
XX BEER; variant; V10I; gene therapy; antisense therapy; fracture;
XX chromosome 17q12-21; bone mineralization; ss.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 48..689
XX FT /*tag= a
XX FT /label= BEER variant V10I
XX FT /product= TGF-beta_binding_protein
XX
XX W0200032773-A1.
XX
XX 08-JUN-2000.
XX
XX 24-NOV-1999; 99WO-US27990.
XX
XX 27-NOV-1998; 98US-0110283.
XX (DARW-) DARWIN DISCOVERY LTD.
XX
XX Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepers BW;
XX Van Ness J, Winkler DG;
XX WPI; 2000-412321/35.
XX P-PSDB; AAY96430.
XX
XX Nucleic acids (1) encoding a transforming growth factor beta binding
XX protein, useful for identifying agents for treating osteopenia,
XX osteoporosis and fractures
XX
XX Claim 1; Page 118-119; 162pp; English.
XX
XX This cDNA encodes a variant human transforming growth factor-beta
XX (TGF-beta) binding protein designated BEER V10I. The encoded protein
XX comprises a substitution of isoleucine for the wild-type valine at
XX residue 10. The cDNA and protein may be used for prevention, treatment
XX and diagnosis of diseases associated with inappropriate BEER expression.
XX For example, they may be used to treat disorders associated with
XX decreased TGF-beta BP expression. The cDNA or vectors may be administered
XX to treat diseases by rectifying mutations or deletions in a patient's
XX genome that affect the activity of BEER by expressing inactive proteins
XX or to supplement the patient's own production of BEER polypeptides. The
XX nucleic acids may be used for recombinant production of BEER, gene
XX therapy, antisense therapy, as probes for diagnostic assays and for
XX functional studies. BEER may be used to raise antibodies and for
XX identification of BEER modulators. BEER antagonists may be used to
XX increase bone mineral content for the treatment of disorders such as
XX osteopenia, osteoporosis, fractures and other disorders associated with
XX low mineral content.
XX
XX Sequence 2301 BP; 569 A; 634 C; 614 G; 484 T; 0 other;
XX
XX Query Match 10.8%; Score 69; DB 21; Length 2301;
XX Best Local Similarity 100.0%; Pred. No. 4.2e-23;
XX Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 475 CTGGTGGCTCGTGAAGTGCAGCGCCTCACCGCTTCCACACACGATCGGAGCTCAAG 534
Db 528 CTGGTGGCTCGTGAAGTGCAGCGCCTCACCGCTTCCACACACGATCGGAGCTCAAG 587
QY 535 GACTTCGGG 543
Db 588 GACTTCGGG 596
```

XX Human; DNA/Cerberus-related protein 6; hDCR6; morphogenic protein;
 KW antagonist; BMP; cell growth; cell differentiation; bone formation;
 KW gene therapy; ss.
 XX Homo sapiens.
 XX Key Location/Qualifiers
 FH 1..1104
 FT /*tag= a
 FT /product= "hDCR6 #1"
 XX WO200055193-A2.
 XX 21-SEP-2000.
 XX 02-MAR-2000; 2003WO-US05537.
 XX 12-MAR-1999; 99US-0124118.
 XX (REG-) REGENERON PHARM INC.
 XX Economides AN;
 XX WPI; 2000-638179/61.
 XX P-PSDB; AAE26105.
 XX Novel isolated, human DNA/Cerberus related protein 6 which include
 PT natural homologue, and polypeptides comprising DCR6 domain and nucleic
 PT acids encoding the proteins which are useful as probes and primers -
 XX Claim 2; Fig 2; 40pp; English.
 XX The present sequence comprises the human DAN/Cerberus-related protein 6
 CC (hDCR6) coding sequence. It was isolated from a genomic DNA clone
 CC following identification using computer-based 'virtual cloning'. hDCR6
 CC is closely related to the DAN and DCR5 proteins, both of which act as
 CC antagonists of morphogenic proteins such as BMP. It is possible that
 CC the hDCR6 gene and protein can be used as immunogens, modulators of cell
 CC function, growth and differentiation, to reduce undesirable bone
 CC formation, to identify DCR6 binding agents, in diagnosis, and in gene
 CC therapy.
 XX Sequence 1104 BP; 226 A; 354 C; 364 G; 160 T; 0 other;
 SQ Query Match 10.8%; Score 69; DB 21; Length 1104;
 Best Local Similarity 100.0%; Pred. No. 4.4e-23;
 Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 475 CTGGTGGCTCTGCAAGTGCAGCGCCTCACCGCTTCCACACAGTCGGAGCTCAAG 534
 Db 943 CTGGTGGCTCTGCAAGTGCAGCGCCTCACCGCTTCCACACAGTCGGAGCTCAAG 1002
 QY 535 GACTTCGGG 543
 Db 1003 GACTTCGGG 1011
 RESULT 12
 ID AAD27576
 XX AAD27576 standard; cDNA; 2271 BP.
 AC AAD27576;
 XX 18-APR-2002 (first entry)
 DT Human osteolevin cDNA.
 DE Human; osteolevin; osteopathic; cytostatic; bone formation; osteoporosis;
 KW Van Buchem-sclerosteosis disease; sclerosteosis; transgenic animal;
 KW Paget's disease; chromosome 17; ss.
 XX Homo sapiens.
 OS

XX Key Location/Qualifiers
 FH 13..634
 FT /*tag= a
 FT /product= "Human osteolevin protein"
 FT sig_peptide 13..69
 FT /*tag= b
 FT mat_peptide 70..631
 FT /*tag= c
 FT /product= "Mature human osteolevin protein"
 XX WO200198491-A2.
 XX 27-DEC-2001.
 XX 15-JUN-2001; 2001WO-E306795.
 XX 19-JUN-2000; 2000EP-0112867.
 XX (HOFF) HOFFMANN LA ROCHE & CO AG F.
 XX (UYIN-) UNIV INSTELLING ANTWERPEN UIA.
 XX Bailemans W, Ebeling M, Foerznier D, Patel N, Van Hul W;
 PI Vickery BH;
 XX WPI; 2002-139789/18.
 XX P-PSDB; AAE17089.
 XX Novel genetic polymorphisms in the Van Buchem-sclerosteosis disease
 PT region that are associated with abnormal bone formation useful for
 PT diagnosis and assessment of osteoporosis or sclerosteosis in humans -
 XX Claim 10; Fig 1B; 70pp; English.
 XX The invention relates to a nucleic acid encoding osteolevin region
 CC polymorphisms. The invention also relates to genetic polymorphisms in
 CC the Van Buchem-sclerosteosis disease region that are associated with
 CC disorders resulting in either net excess bone formation or insufficient
 CC bone formation in humans. Osteolevin DNA is useful for screening for
 CC osteolevin polymorphisms associated with abnormal bone formation in a
 CC subject and for the presence of a heritably linked form of abnormal bone
 CC formation in a subject, by determining the presence of a polymorphism in
 CC the osteolevin nucleic acid sequence obtained from the subject.
 CC Osteolevin protein is useful for treating diseases associated with
 CC abnormal bone formation, such as sclerosteosis, Van Buchem's disease and
 CC Paget's disease. Nucleic acids which encode genes in the osteolevin
 CC region or their modified forms can also be used to generate either
 CC transgenic animals or knockout animals which are useful in the screening
 CC and development of therapeutically useful reagents. Osteolevin proteins
 CC are useful in pharmacological characterisation of novel modulators of the
 CC activity of protein and protein complexes. Human osteolevin gene is
 CC located on chromosome 17. The present sequence is human osteolevin cDNA.
 XX Sequence 2271 BP; 573 A; 623 C; 599 G; 476 T; 0 other;
 SQ Query Match 10.8%; Score 69; DB 24; Length 2271;
 Best Local Similarity 100.0%; Pred. No. 4.2e-23;
 Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 475 CTGGTGGCTCTGCAAGTGCAGCGCCTCACCGCTTCCACACAGTCGGAGCTCAAG 534
 Db 493 CTGGTGGCTCTGCAAGTGCAGCGCCTCACCGCTTCCACACAGTCGGAGCTCAAG 552
 QY 535 GACTTCGGG 543
 Db 553 GACTTCGGG 561
 RESULT 13
 ID AAA29055
 XX AAA29055 standard; cDNA; 2301 BP.
 XX AC AAA29055;

```
FT CDS 1..642
FT /*tag= a
FT /product= "hDCR6 #2"
FT /partial
FT
XX WO200055193-A2.
XX
XX 21-SEP-2000.
XX
XX 02-MAR-2000; 2000WO-US05537.
XX
XX 12-MAR-1999; 99US-0124118.
XX
XX (REGE-) REGENERON PHARM INC.
XX
XX Economides AN;
XX
XX WPI; 2000-638179/61.
XX P-PSDB; AAB26106.
XX
XX Novel isolated, human DNA/Cerberus related protein 6 which include
XX PT natural homologue, and polypeptides comprising DCR6 domain and nucleic
XX PT acids encoding the proteins which are useful as probes and primers -
XX
XX Claim 2; Fig 3; 40pp; English.
XX
XX The present sequence comprises exons 1 and 4 of the human
XX CC DAN/Cerberus-related protein 6 (hDCR6) coding sequence. It was isolated
XX CC from a human kidney cDNA library containing exons 1 and 4 of the
XX CC sequence. hDCR6 is closely related to the DAN and DCR5 proteins, both of
XX CC which act as antagonists of morphogenic proteins such as BMP. It is
XX CC possible that the hDCR6 gene and protein can be used as immunogens,
XX CC modulators of cell function, growth and differentiation, to reduce
XX CC undesirable bone formation, to identify DCR6 binding agents, in
XX CC diagnosis, and in gene therapy.
XX
XX Sequence 642 BP; 115 A; 230 C; 211 G; 86 T; 0 other;
XX
XX Query Match 10.8%; Score 69; DB 21; Length 642;
XX Best Local Similarity 100.0%; Pred. No. 4.6e-23;
XX Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 475 CTGGTGGCTCGTGAAGTGCAGCGCCTCACCGCTTCCACACAGTCGAGCTCAAG 534
XX |||||||
XX Db 481 CTGGTGGCTCGTGAAGTGCAGCGCCTCACCGCTTCCACACAGTCGAGCTCAAG 540
XX |||||||
XX QY 535 GACTTCGGG 543
XX |||||||
XX Db 541 GACTTCGGG 549
XX
XX RESULT 10
XX ABA94293
XX ID ABA94293 standard; cDNA; 759 BP.
XX
XX AC ABA94293;
XX
XX DT 26-MAR-2002 (first entry)
XX
XX DE Human cloaked-2 polypeptide encoding cDNA.
XX
XX Cloaked-2; cysteine knot motif; nephrotropic; cardiant; immunomodulator;
XX KW hepatotropic; antiinflammatory; antithyroid; cytostatic; neuroprotective;
XX KW antianemic; hypotensive; antiarrhythmic; antiarteriosclerotic; muscular;
XX KW antidiabetic; anorectic; gene therapy; cell therapy; antisense therapy;
XX KW human; ss.
XX
XX OS Homo sapiens.
XX
XX XX Location/Qualifiers
XX Key 37..678
XX FT /*tag= a
XX FT /product= "cloaked-2 polypeptide"
XX
```

```
FT sig_peptide 37..105
FT /*tag= b
FT mat_peptide 106..759
FT /*tag= c
FT /note= "see ABB07207"
XX
XX WO200192308-A2.
XX
XX 06-DEC-2001.
XX
XX 29-MAY-2001; 2001WO-US17478.
XX
XX 01-JUN-2000; 2000US-208550P.
XX
XX 04-AUG-2000; 2000US-223542P.
XX
XX (AMGE-) AMGEN INC.
XX
XX Paszty CJ, Gao Y;
XX
XX WPI; 2002-114325/15.
XX P-PSDB; ABB07207, ABB07209.
XX
XX New human and mouse cysteine-knot polypeptide designated as Cloaked-2,
XX PT for treating or preventing kidney, heart (e.g. myocardial infarction)
XX PT or liver (e.g. hepatitis) diseases -
XX
XX Claim 1; Fig 1; 170pp; English.
XX
XX The invention relates to polypeptides comprising a cysteine knot motif
XX CC and designated as Cloaked-2, derived from human and mouse. The cloaked-2
XX CC polypeptides can be expressed by standard recombinant methodology. The
XX CC cloaked-2 polynucleotides are useful in gene therapy and antisense
XX CC therapy. The cloaked-2 polypeptides and polynucleotides are useful for
XX CC treating, preventing, ameliorating or detecting diseases and disorders of
XX CC the kidney (e.g. anemia, hypertension or low blood pressure), heart (e.g.
XX CC cardiac hypertrophy, congestive heart failure, myocardial infarction,
XX CC arrhythmias, atherosclerosis, hypertension or low blood pressure),
XX CC skeletal muscle (e.g. muscular dystrophy or cachexia), placenta (e.g.
XX CC congenital abnormalities or miscarriage), liver (e.g. hepatitis or
XX CC cirrhosis), pancreas (e.g. diabetes or pancreatitis), thyroid (e.g.
XX CC Grave's disease or myxedema) or adrenal cortex (e.g. Cushing's disease
XX CC or Addison's disease), homeostasis or metabolic diseases (e.g. obesity,
XX CC cancer or myopathies), infections, or autoimmune diseases. Selective
XX CC binding agents may be used to modulate the biological activities of
XX CC Cloaked-2 polypeptides or to detect Cloaked-2 polypeptide levels in a
XX CC sample. Transgenic non-human animals are useful for drug candidate
XX CC screening. The present sequence represents a cDNA encoding the human
XX CC cloaked-2 polypeptide.
XX
XX Sequence 759 BP; 125 A; 282 C; 244 G; 108 T; 0 other;
XX
XX Query Match 10.8%; Score 69; DB 24; Length 759;
XX Best Local Similarity 100.0%; Pred. No. 4.5e-23;
XX Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 475 CTGGTGGCTCGTGAAGTGCAGCGCCTCACCGCTTCCACACAGTCGAGCTCAAG 534
XX |||||||
XX Db 517 CTGGTGGCTCGTGAAGTGCAGCGCCTCACCGCTTCCACACAGTCGAGCTCAAG 576
XX |||||||
XX
XX QY 535 GACTTCGGG 543
XX |||||||
XX Db 577 GACTTCGGG 585
XX
XX RESULT 11
XX AAA94050
XX ID AAA94050 standard; DNA; 1104 BP.
XX
XX AC AAA94050;
XX
XX XX Location/Qualifiers
XX DT 15-JAN-2001 (first entry)
XX
XX DE Human DAN/Cerberus-related protein 6 (hDCR6) coding sequence #1.
```


PT analyzing gene expression in human bone marrow -

XX Example 4; SEQ ID NO: 20066; 658pp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC bone marrow. They can be used to measure gene expression in bone marrow
 CC samples, which may enable the improved diagnosis and treatment of cancers
 CC such as lymphoma, leukaemia and myeloma. The present sequence is one of
 CC the probes of the invention.

XX Sequence 422 BP; 53 A; 145 C; 159 G; 65 T; 0 other;

Query Match 10.8%; Score 69; DB 23; Length 422;

Best Local Similarity 100.0%; Pred. No. 4.7e-23;

Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 475 CTGGTGGCTCTGCAAGTGCAGCGCTTCCACCGCTTCCACACAGTGGAGCTCAAG 534

Db 162 CTGGTGGCTCTGCAAGTGCAGCGCTTCCACCGCTTCCACACAGTGGAGCTCAAG 103

Qy 535 GACTTCGGG 543

Db 102 GACTTCGGG 94

RESULT 8

ABS19777/C

ID ABS19777 standard; DNA; 422 BP.

XX ABS19777;

XX ABS19777;

XX 19-AUG-2002 (first entry)

XX Human genome-derived single exon probe ORF from lung SEQ ID No 19768.

XX Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
 KW chronic obstructive pulmonary disease; interstitial lung disease;
 KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
 KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
 KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
 KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
 KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
 KW primary ciliary dyskinesia; pulmonary hypertension;
 KW hyaline membrane disease; open reading frame; ORF.

XX Homo sapiens.

XX WO200186003-A2.

XX 15-NOV-2001.

XX 30-JAN-2001; 2001WO-US00665.

XX 04-FEB-2000; 2000US-180312P.

XX 26-MAY-2000; 2000US-207456P.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0632366.

XX 21-SEP-2000; 2000US-234687P.

XX 27-SEP-2000; 2000US-236359P.

XX 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2002-114183/15.

XX Spatially-addressable set of single exon nucleic acid probes, used to
 PT measure gene expression in human lung samples -

XX Claim 4; SEQ ID No 19768; 634pp; English.

CC The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human lung comprising single exon nucleic acid probes having one of
 CC 12614 nucleic acid sequences mentioned in the specification, or their
 CC complements or the 12617 open reading frames derived from the 12614
 CC probes. Also included are a microarray comprising the novel set of
 CC probes; the novel set of probes which hybridise at high stringency to a
 CC nucleic acid expressed in the human lung; measuring gene expression in a
 CC sample derived from human lung, comprising (a) contacting the array with
 CC a collection of detectably labeled nucleic acids derived from human lung
 CC mRNA, and (b) measuring the label detectably bound to each probe of
 CC the array; identifying exons in a eukaryotic genome, comprising
 CC (a) algorithmically predicting at least one exon from genomic sequences
 CC of the eukaryote; and (b) detecting specific hybridisation of detectably
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
 CC having a fragment identical to the predicted exon, the probe is included
 CC in the above mentioned microarray; assigning exons to a single gene,
 CC comprising (a) identifying exons from genomic sequence by the method
 CC above and (b) measuring the expression of each of the exons in several
 CC tissues and/or cell types using hybridisation to a single exon
 CC microarrays having a probe with the exon, where a common pattern of
 CC expression of the exons in the tissues and/or cell types indicates that
 CC the exons should be assigned to a single gene; a peptide comprising one
 CC of 12011 sequences, mentioned in the specification, or encoded by the
 CC probes/open reading frames (ORF). The probes are used for gene
 CC expression analysis, and for identifying exons in a gene, particularly
 CC using human lung derived mRNA and for the study of lung diseases
 CC such as asthma, lung cancer, chronic obstructive pulmonary disease
 CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
 CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
 CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
 CC haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
 CC pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic
 CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
 CC and hyaline membrane disease. The present sequence is a single exon
 CC probe open reading frame of the invention.

CC Note: The sequence data for this patent did not form part

CC of the printed specification, but was obtained in electronic

CC format directly from WIPO at

CC ftp.wipo.int/pub/publ:shed_pct_sequences.

XX Sequence 422 BP; 53 A; 145 C; 159 G; 65 T; 0 other;

Query Match 10.8%; Score 69; DB 24; Length 422;

Best Local Similarity 100.0%; Pred. No. 4.7e-23;

Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 475 CTGGTGGCTCTGCAAGTGCAGCGCTTCCACCGCTTCCACACAGTGGAGCTCAAG 534

Db 162 CTGGTGGCTCTGCAAGTGCAGCGCTTCCACCGCTTCCACACAGTGGAGCTCAAG 103

Qy 535 GACTTCGGG 543

Db 102 GACTTCGGG 94

RESULT 9

AAA94051

ID AAA94051 standard; DNA; 642 BP.

XX AAA94051;

XX 15-JAN-2001 (first entry)

XX Human DAN/Cerberus-related protein 6 (hDCR6) cDNA exons 1 and 4.

XX Human; DNA/Cerberus-related protein 6; hDCR6; morphogenic protein;
 KW antagonist; BMP; cell growth; cell differentiation; bone formation;
 KW gene therapy; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

Db 280 CTGGTGGCCTGCAAGTGCAGCGCCTCACCCTTCCACACCAAGTCGAGCTCAAG 221

QY 535 GACTTCGGG 543

Db 220 GACTTCGGG 212

RESULT 6

ID ABS07272/c

XX ABS07272 standard; DNA; 392 BP.

AC ABS07272;

XX 19-AUG-2002 (first entry)

XX Human genome-derived single exon probe from lung SEQ ID No 7263.

DE Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;

XX Chronic obstructive pulmonary disease; interstitial lung disease;

KW familial idiopathic pulmonary fibrosis; neurofibromatosis;

KW tuberosus sclerosis; Gaucher's disease; Niemann-Pick disease;

KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;

KW pulmonary histiocytosis; lymphangioma; lymphoma; Karagener syndrome;

KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;

KW primary ciliary dyskinesia; pulmonary hypertension;

KW hyaline membrane disease.

XX Homo sapiens.

XX WO200186003-A2.

PN 15-NOV-2001.

XX 30-JAN-2001; 2001WO-US00665.

XX 04-FEB-2000; 2000US-180312P.

PR 26-MAY-2000; 2000US-207456P.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-234687P.

PR 27-SEP-2000; 2000US-236359P.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PA Penn SG, Hanzel DK, Chen W, Rank DR;

PI WPI; 2002-114183/15.

DR Spatially-addressable set of single exon nucleic acid probes, used to

XX measure gene expression in human lung samples -

PS Claim 1; SEQ ID No 7263; 634pp; English.

XX The invention relates to a spatially-addressable set of single exon

CC nucleic acid probes for measuring gene expression in a sample derived

CC from human lung comprising single exon nucleic acid probes having one of

CC 12614 nucleic acid sequences mentioned in the specification, or their

CC complements or the 12387 open reading frames derived from the 12614

CC probes. Also included are a microarray comprising the novel set of

CC probes; the novel set of probes which hybridise at high stringency to a

CC nucleic acid expressed in the human lung; measuring gene expression in a

CC sample derived from human lung, comprising (a) contacting the array with

CC a collection of detectably labeled nucleic acids derived from human lung

CC mRNA, and (b) measuring the label detectably bound to each probe of

CC the array; identifying exons in a eukaryotic genome, comprising

CC (a) algorithmically predicting at least one exon from genomic sequences

CC of the eukaryote; and (b) detecting specific hybridisation of detectably

CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,

CC having a fragment identical to the predicted exon, the probe is included

CC in the above mentioned microarray; assigning exons to a single gene,

CC comprising (a) identifying exons from genomic sequence by the method

CC above and (b) measuring the expression of each of the exons in several

CC tissues and/or cell types using hybridisation to a single exon

CC microarrays having a probe with the exon, where a common pattern of

CC expression of the exons in the tissues and/or cell types indicates that

CC the exons should be assigned to a single gene; a peptide comprising one

CC of 12011 sequences, mentioned in the specification, or encoded by the

CC probes/open reading frames (ORF). The probes are used for gene

CC expression analysis, and for identifying exons in a gene, particularly

CC using human lung derived mRNA and for the study of lung diseases

CC such as asthma, lung cancer, chronic obstructive pulmonary disease

CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary

CC fibrosis, neurofibromatosis, tuberosus sclerosis, Gaucher's disease,

CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary

CC haemosiderosis, pulmonary histiocytosis, lymphangioma, lymphoma,

CC pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic

CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension

CC and hyaline membrane disease. The present sequence is a single exon

CC probe of the invention.

CC Note: The sequence data for this patent did not form part

CC of the printed specification, but was obtained in electronic

CC format directly from WIPO at

CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 392 BP; 60 A; 123 C; 149 G; 60 T; 0 other;

QY Query Match 10.8%; Score 69; DB 24; Length 392;

Db Best Local Similarity 100.0%; Pred. No. 4.7e-23;

XX Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 475 CTGGTGGCCTGCAAGTGCAGCGCCTCACCCTTCCACACCAAGTCGAGCTCAAG 534

Db 280 CTGGTGGCCTGCAAGTGCAGCGCCTCACCCTTCCACACCAAGTCGAGCTCAAG 221

QY 535 GACTTCGGG 543

Db 220 GACTTCGGG 212

RESULT 7

AAK45509/c

ID AAK45509 standard; DNA; 422 BP.

XX AAK45509;

AC AAK45509;

XX 06-NOV-2001 (first entry)

DT Human bone marrow expressed single exon probe SEQ ID NO: 20066.

XX Human; bone marrow expressed exon; gene expression analysis; probe;

KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.

XX Homo sapiens.

XX WO200157276-A2.

PN 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00668.

XX 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PA Penn SG, Hanzel DK, Chen W, Rank DR;

PI WPI; 2001-488900/53.

XX Human genome-derived single exon nucleic acid probes useful for


```
CC associated with low mineral content.
XX
SQ Sequence 638 BP; 114 A; 224 C; 209 G; 91 T; 0 other;

Query Match          92.0%; Score 585; DB 21; Length 638;
Best Local Similarity 99.8%; Pred. No. 2.3e-269;
Matches 635; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGCAGCCCTCACTAGCCCGCTCATCTGCTACTTGTGTGACAGCTGCTTCTGTGCT 60
DB 1 ATGCAGCCCTCACTAGCCCGCTCATCTGCTACTTGTGTGACAGCTGCTTCTGTGCT 60

QY 61 GTGAGGGCCAGGGGTGGCAAGCTTTCAGGAATGATGACCAAGAGTCTATCCAGGGCTT 120
DB 61 GTGAGGGCCAGGGGTGGCAAGCTTTCAGGAATGATGACCAAGAGTCTATCCAGGGCTT 120

QY 121 GGAGGTACCCCGAGCTCTCTGTAGAACCAACGACCATGACCGGGGGAGATGA 180
DB 121 GGAGGTACCCCGAGCTCTCTGTAGAACCAACGACCATGACCGGGGGAGATGA 180

QY 181 GGCAGACCTCCCAACCATCCCTATGACGCGCAAGATGTGTCCAGTACAGCTGCCGCGAG 240
DB 181 GGCAGACCTCCCAACCATCCCTATGACGCGCAAGATGTGTCCAGTACAGCTGCCGCGAG 240

QY 241 CTGCACTACACCGCTTCTGTAGACAGCGCCCATGCGGAGCGCCAAAGCGGTCAACGAG 300
DB 241 CTGCACTACACCGCTTCTGTAGACAGCGCCCATGCGGAGCGCCAAAGCGGTCAACGAG 300

QY 301 TTGCTGTGCTCCGCGCAGTGCAGCGCCCGCGCGCTGCTGCCCAACGCCATCGGGCGGTG 360
DB 301 TTGCTGTGCTCCGCGCAGTGCAGCGCCCGCGCGCTGCTGCCCAACGCCATCGGGCGGTG 360

QY 361 AAGTGTGGCGCCGCAACGACCGGATTTCCGCTGCATCCCGGATCGCTACCGCGCGAG 420
DB 361 AAGTGTGGCGCCGCAACGACCGGATTTCCGCTGCATCCCGGATCGCTACCGCGCGAG 420

QY 421 CGGCTGAGCTGTGTGCTCCCGGGCGCGCGCGCTGCGCGAAGTGTGCTGTGTTG 480
DB 421 CGGCTGAGCTGTGTGCTCCCGGGCGCGCGCGCTGCGCGAAGTGTGCTGTGTTG 480

QY 481 GCCTGTGCAAGTGAAGCGCTCACCGCTTCCCAACCATGTCGAGCTCAAGACTTC 540
DB 481 GCCTGTGCAAGTGAAGCGCTCACCGCTTCCCAACCATGTCGAGCTCAAGACTTC 540

QY 541 GGGCCGAGACCGCGCGCGCGCAAGAGGTGCGCAAGCGCGCGCGCGCGCGAGCC 600
DB 541 GGGCCGAGACCGCGCGCGCGCAAGAGGTGCGCAAGCGCGCGCGCGCGCGAGCC 600

QY 601 AAGCCCAACGCGGAGCTGGAGACGCTACTAG 636
DB 601 AAGCCCAACGCGGAGCTGGAGACGCTACTAG 636

RESULT 3
AA29063
ID AAA29063 standard; DNA; 35828 BP.
XX
AC AAA29063;
XX
DT 12-SEP-2000 (first entry)
XX
DE Murine TGF-beta binding protein (BEER) genomic MluI-AviII DNA fragment.
XX Osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
KW BEER; gene therapy; antisense therapy; fracture; bone mineralization;
KW transgenic mouse; over-expression; ds.
XX
OS Mus musculus.
XX
FH Key Location/Qualifiers
FT CDS 18317..21464
FT FT /*tag= a
FT 18317..18530
FT exon
```

PF 29-MAY-2001; 2001WO-US17478.
 XX
 PR 01-JUN-2000; 2000US-208550P.
 PR 04-AUG-2000; 2000US-223542P.
 XX
 PA (AMGE-) AMGEN INC.
 PI Paszty CJ, Gao Y;
 XX
 XX WPI: 2002-114325/15.
 DR P-PSDB; ABB07208, ABB07210.
 XX
 PT New human and mouse cysteine-knot polypeptide designated as Cloaked-2,
 PT for treating or preventing kidney, heart (e.g. myocardial infarction)
 PT or liver (e.g. hepatitis) diseases
 XX
 PS Claim 1; Fig 2; 170pp; English.
 XX
 CC The invention relates to polypeptides comprising a cysteine knot motif
 CC and designated as Cloaked-2, derived from human and mouse. The cloaked-2
 CC polypeptides can be expressed by standard recombinant methodology. The
 CC cloaked-2 polynucleotides are useful in gene therapy and antisense
 CC therapy. The cloaked-2 polypeptides and polynucleotides are useful for
 CC treating, preventing, ameliorating or detecting diseases and disorders of
 CC the kidney (e.g. anemia, hypertension or low blood pressure), heart (e.g.
 CC cardiac hypertrophy, congestive heart failure, myocardial infarction,
 CC arrhythmias, atherosclerosis, hypertension or low blood pressure),
 CC skeletal muscle (e.g. muscular dystrophy or cachexia), placenta (e.g.
 CC congenital abnormalities or miscarriage), liver (e.g. hepatitis or
 CC cirrhosis), pancreas (e.g. diabetes or pancreatitis), thyroid (e.g.
 CC Grave's disease or myxedema) or adrenal cortex (e.g. Cushing's disease
 CC or Addison's disease), homeostasis or metabolic diseases (e.g. obesity,
 CC cancer or myopathies), infections, or autoimmune diseases. Selective
 CC binding agents may be used to modulate the biological activities of
 CC Cloaked-2 polypeptides or to detect Cloaked-2 polypeptide levels in a
 CC sample. Transgenic non-human animals are useful for drug candidate
 CC screening. The present sequence represents a cDNA encoding the mouse
 CC cloaked-2 polypeptide.
 XX
 SQ Sequence 636 BP; 114 A; 224 C; 207 G; 91 T; 0 other;

Query Match 100.0%; Score 636; DB 24; Length 636;
 Best Local Similarity 100.0%; Pred. No. 1e-293;
 Matches 636; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCAGCCCTACTAGCCCGGCTCATCTGCTACTTGTGACGCTGCTCTGCTCT 60
 Db
 QY 61 GTGGAGGGCCAGGGGTGGCAAGCCTTCAGGAATGATGCCACAGAGGTATCCCGGGCTT 120
 Db 61 GTGGAGGGCCAGGGGTGGCAAGCCTTCAGGAATGATGCCACAGAGGTATCCCGGGCTT 120
 QY 121 GGAGAGTACCCGAGGCTCTCTGAGAACACACAGACCATGACCGGGGAGAGATGGA 180
 Db 121 GGAGAGTACCCGAGGCTCTCTGAGAACACACAGACCATGACCGGGGAGAGATGGA 180
 QY 181 GGCAGACCTCCACCATCTCCATGACGCGCAAGATGTCTGAGTACAGTGCAGCGGAG 240
 Db 181 GGCAGACCTCCACCATCTCCATGACGCGCAAGATGTCTGAGTACAGTGCAGCGGAG 240
 QY 241 CTGCACTACACCGCTTCTTGACAGACGCGCCCATGCCGAGCGCCCAAGCCGCTCACCAG 300
 Db 241 CTGCACTACACCGCTTCTTGACAGACGCGCCCATGCCGAGCGCCCAAGCCGCTCACCAG 300
 QY 301 TTGGTGTGCTCGGCGAGTGCAGGCGCGCGGCTGTGTCGCAAGCCATCGGGCGGTG 360
 Db 301 TTGGTGTGCTCGGCGAGTGCAGGCGCGCGGCTGTGTCGCAAGCCATCGGGCGGTG 360
 QY 361 AAGTGTGCGCGCCCAAGACGAGCGGATTTCCGCTGCATCCCGGATCGTACGCGCGGAG 420
 Db 361 AAGTGTGCGCGCCCAAGACGAGCGGATTTCCGCTGCATCCCGGATCGTACGCGCGGAG 420

QY 421 CGGGTGCAGCTGTGTGCTCCCGGGGGCGCGCGCTCGCGCAAGGTGCGTCTGGTG 480
 Db
 QY 421 CGGGTGCAGCTGTGTGCTCCCGGGGGCGCGCGCTCGCGCAAGGTGCGTCTGGTG 480
 Db
 QY 481 GCCTGTGCAAGTGCAGCGCTTCCACCGCTTCCACCAACAGTCCGAGCTCAAGGACTTC 540
 Db 481 GCCTGTGCAAGTGCAGCGCTTCCACCGCTTCCACCAACAGTCCGAGCTCAAGGACTTC 540
 QY 541 GGGCGGAGACCGCGCGCGCGAGAGGTCGCAAGCGCGCGCGCGCGCGCGCGGAGCC 600
 Db 541 GGGCGGAGACCGCGCGCGCGAGAGGTCGCAAGCGCGCGCGCGCGCGCGCGGAGCC 600
 QY 601 AAAGCCCAACAGCGGAGCTGGAGAACGCTTACTAG 636
 Db 601 AAAGCCCAACAGCGGAGCTGGAGAACGCTTACTAG 636

RESULT 2
 AAA29058
 ID AAA29058 standard; cDNA; 638 BP.
 XX
 AC AAA29058;
 XX
 DT 12-SEP-2000 (first entry)
 XX
 DE Murine TGF-beta binding protein (BEER) cDNA.
 XX
 KW osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
 KW BEER; gene therapy; antisense therapy; fracture; bone mineralization; ss.
 XX
 OS Mus musculus.
 PH Key Location/Qualifiers
 FT CDS 1..636
 FT /*tag: a
 FT /product= TGF-beta_binding_protein
 XX
 PN WO200032773-A1.
 PD 08-JUN-2000.
 XX
 PF 24-NOV-1999; 99WO-US27990.
 XX
 PR 27-NOV-1998; 98US-0:10283.
 XX
 XX (DARW-) DARWIN DISCOVERY LTD.
 XX
 PI Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepker BW;
 PI Van Ness J, Winkler DG;
 XX
 XX WPI: 2000-412321/35.
 DR P-PSDB; AAY96432.
 XX
 PT Nucleic acids (I) encoding a transforming growth factor beta binding
 PT protein, useful for identifying agents for treating osteopenia,
 PT osteoporosis and fractures
 XX
 PS Claim 1; Page 123; 16pp; English.
 XX
 CC This cDNA encodes a murine transforming growth factor-beta (TGF-beta)
 CC binding protein designated mBEER. The cDNA and protein may be used for
 CC prevention, treatment and diagnosis of diseases associated with
 CC inappropriate BEER expression. For example, they may be used to treat
 CC disorders associated with decreased TGF-beta BP expression. The cDNA or
 CC vectors may be administered to treat diseases by rectifying mutations or
 CC deletions in a patient's genome that affect the activity of BEER by
 CC expressing inactive proteins or to supplement the patients own production
 CC of BEER polypeptides. The nucleic acids may be used for recombinant
 CC production of BEER, gene therapy, antisense therapy, as probes for
 CC diagnostic assays and for functional studies. BEER may be used to raise
 CC antibodies and for identification of BEER modulators. BEER antagonists
 CC may be used to increase bone mineral content for the treatment of
 CC disorders such as osteopenia, osteoporosis, fractures and other disorders

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 28, 2003, 23:23:29 ; Search time 199.234 Seconds
(without alignments)
7188.873 Million cell updates/sec

Title: US-09-867-274-3
Perfect score: 636
Sequence: 1 atgcagccctactagcccc.....agctggagaaagcctactag 636

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 2185239 seqs, 1125999159 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq 101002.*
1: /SID22/gcgdata/geneseq/geneseqn-embl/NA1980.DAT.*
2: /SID22/gcgdata/geneseq/geneseqn-embl/NA1981.DAT.*
3: /SID22/gcgdata/geneseq/geneseqn-embl/NA1982.DAT.*
4: /SID22/gcgdata/geneseq/geneseqn-embl/NA1983.DAT.*
5: /SID22/gcgdata/geneseq/geneseqn-embl/NA1984.DAT.*
6: /SID22/gcgdata/geneseq/geneseqn-embl/NA1985.DAT.*
7: /SID22/gcgdata/geneseq/geneseqn-embl/NA1986.DAT.*
8: /SID22/gcgdata/geneseq/geneseqn-embl/NA1987.DAT.*
9: /SID22/gcgdata/geneseq/geneseqn-embl/NA1988.DAT.*
10: /SID22/gcgdata/geneseq/geneseqn-embl/NA1989.DAT.*
11: /SID22/gcgdata/geneseq/geneseqn-embl/NA1990.DAT.*
12: /SID22/gcgdata/geneseq/geneseqn-embl/NA1991.DAT.*
13: /SID22/gcgdata/geneseq/geneseqn-embl/NA1992.DAT.*
14: /SID22/gcgdata/geneseq/geneseqn-embl/NA1993.DAT.*
15: /SID22/gcgdata/geneseq/geneseqn-embl/NA1994.DAT.*
16: /SID22/gcgdata/geneseq/geneseqn-embl/NA1995.DAT.*
17: /SID22/gcgdata/geneseq/geneseqn-embl/NA1996.DAT.*
18: /SID22/gcgdata/geneseq/geneseqn-embl/NA1997.DAT.*
19: /SID22/gcgdata/geneseq/geneseqn-embl/NA1998.DAT.*
20: /SID22/gcgdata/geneseq/geneseqn-embl/NA1999.DAT.*
21: /SID22/gcgdata/geneseq/geneseqn-embl/NA2000.DAT.*
22: /SID22/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT.*
23: /SID22/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT.*
24: /SID22/gcgdata/geneseq/geneseqn-embl/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	636	100.0	636	24 ABA94294	Mouse cloaked-2 po
2	585	92.0	638	21 AAA29058	Murine TGF-beta bi
3	424	66.7	35828	21 AAA29063	Murine TGF-beta bi
4	104	16.4	674	21 AAA29059	Rat TGF-beta bindi
5	69	10.8	392	22 AAK32487	Human bone marrow
6	69	10.8	392	22 ABS07272	Human genome-deriv
7	69	10.8	422	22 AAK45509	Human bone marrow
8	69	10.8	422	22 ABS19777	Human genome-deriv
9	69	10.8	642	21 AAA94051	Human DAN/Cerberus

10	69	10.8	759	24 ABA94293	Human cloaked-2 po
11	69	10.8	1104	21 AAA94050	Human DAN/Cerberus
12	69	10.8	2271	24 AAD27576	Human osteolevin c
13	69	10.8	2301	21 AAA29055	Human TGF-beta bin
14	69	10.8	2301	21 AAA29056	Human TGF-beta bin
15	69	10.8	2301	21 AAA29061	Mutant human TGF-b
16	69	10.8	2301	21 AAA29062	Human TGF-beta bin
17	69	10.8	2329	22 AAA91023	Human secreted pro
18	69	10.8	2329	24 ABK69992	cDNA encoding huma
19	69	10.8	5680	21 AAA94049	Human DAN/Cerberus
20	69	10.8	7099	24 AAD27577	Human osteolevin g
21	69	10.8	9301	21 AAA29064	Human TGF-beta bin
22	68	10.7	642	21 AAA29057	Vervet TGF-beta bi
23	32	5.0	532	21 AAA29060	Bovine TGF-beta bi
24	26	4.1	41	24 ABA94308	Mouse cloaked-2 cd
25	25	3.9	25	24 ABA94302	Mouse cloaked-2 cd
26	25	3.9	45	24 ABA94307	Mouse cloaked-2 cd
27	24	3.8	24	24 ABA94300	Mouse cloaked-2 cd
28	24	3.8	30	21 AAA94047	Human hDCE6 exon 4
29	23	3.6	555	24 ABQ35254	Oligonucleotide fo
30	23	3.6	555	24 ABQ35255	Oligonucleotide fo
31	21	3.3	21	24 ABA94303	Mouse cloaked-2 cd
32	21	3.3	21	24 ABA94305	Mouse cloaked-2 cd
33	21	3.3	27	22 AAA91035	PCR primer for Hum
34	21	3.3	29	24 ABA94297	Mouse cloaked-2 cd
35	21	3.3	38	21 AAA94048	Human hDCE6 exon 4
36	21	3.3	577	22 ABA64066	Human foetal liver
37	21	3.3	577	22 ABA31229	Probe #9695 for ge
38	21	3.3	577	22 AAK12550	Human brain expres
39	21	3.3	577	22 AAK38268	Human bone marrow
40	21	3.3	577	22 AAL19061	Probe #8994 for ge
41	21	3.3	577	22 AAI44210	Probe #12896 used
42	21	3.3	577	24 ABS12313	Human genome-deriv
43	20	3.1	24	24 ABA94298	Mouse cloaked-2 cd
44	20	3.1	196	21 AAC05741	Human secreted pro
45	20	3.1	198	20 AAX40842	Human secreted pro

ALIGNMENTS

RESULT 1
ABA94294
ID ABA94294 standard; cDNA; 636 BP.

AC ABA94294;

DT 26-MAR-2002 (first entry)

DE Mouse cloaked-2 polypeptide encoding cDNA.

XX Cloaked-2; cysteine knot motif; nephrotropic; cardiant; immunomodulator;
KW hepatotropic; antinflammatory; antithyroid; cytostatic; neuroprotective;
KW antianemic; hypotensive; antiarrhythmic; antiarteriosclerotic; muscular;
KW antidiabetic; anorectic; gene therapy; cell therapy; antisense therapy;
mouse; ss.

OS Mus musculus.

XX Key Location/Qualifiers

FT CDS 1..636

FT FT /*tag= a /product= "cloaked-2 polypeptide"

FT FT 1..69

FT FT /*tag= b

FT FT 70..633

FT FT /*tag= c

FT FT /note= "see ABB07207"

XX WO200192308-A2.

XX 06-DEC-2001.

FD

Mon Mar 31 09:44:28 2003

us-09-867-274-3.oligo.rge

```

VERSION      AF326739.1  GI:13161019
KEYWORDS
SOURCE
ORGANISM      Homo sapiens.
REFERENCE     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS      Brunkow, M.E., Gardner, J.C., Van Ness, J., Paepker, B.W.,
              Kovacevich, B.R., Proll, S., Skonier, J.E., Zhao, L., Sabo, P.J.,
              Fu, Y.H., Alisch, R.S., Gillett, L., Colbert, T., Tacconi, P., Galas, D.,
              Hamersma, H., Beighton, P. and Mulligan, J.T.
              Bone dysplasia sclerosteosis results from loss of the SOST gene
              product, a novel cysteine knot-containing protein
              Am. J. Hum. Genet. 68 (3), 577-589 (2001)
TITLE
JOURNAL      MEDLINE
PUBMED      21090529
REFERENCE     11179006
AUTHORS      2 (bases 1 to 2323)
              Brunkow, M.E., Gardner, J.C., Van Ness, J., Paepker, B.W.,
              Kovacevich, B.R., Proll, S., Skonier, J.E., Zhao, L., Sabo, P.J.,
              Fu, Y.H., Alisch, R.S., Gillett, L., Colbert, T., Tacconi, P., Galas, D.,
              Hamersma, H., Beighton, P. and Mulligan, J.T.
              Direct Submission
              Submitted (07-DEC-2000) Genomics, Celltech Chiroscience Inc., 1631
              220th St. SE, Bothell, WA 98021, USA
TITLE
JOURNAL      Location/Qualifiers
FEATURES
source
1..2323
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="17"
/map="17q21"
1..47
48..689
/codon_start=1
/product="sclerostin"
/protein_id="AAK13454.1"
/db_xref="GI:13161020"
/translation="MOLPLALCIIVLHVHTAPRVVEGQWQAFKNDATETIIPELGEYF
EPPELENNKTMRANGRRPHHPFTKDVSEYSCRELHFTYVYVDGPCRSAPKPVTE
LVCSGQCGPARLLPNAIGKRWKRRPSGDPFRCIPDRYRAORVQLLCPGGEAPRARKVR
LVASCKCKRLTRFHNSQLKDFGTAAARPKGRKPRFRARSAKANQAELFNAY"
690..2323
3'UTR      590 a 635 c 614 g 484 t
BASE COUNT
ORIGIN
Query Match      10.8%; Score 69; DB 9; Length 2323;
Best Local Similarity 100.0%; Pred. No. 3.6e-23;
Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 475 CTGGTGGCCTCGTCAAGTCAAGCGCTCACCGCTTCCACCAACCGAGTCGGAGCTCAAG 534
Db 528 CTGGTGGCCTCGTCAAGTCAAGCGCTCACCGCTTCCACCAACCGAGTCGGAGCTCAAG 587
Qy 535 GACTTCGGG 543
Db 588 GACTTCGGG 596

```

Search completed: March 29, 2003, 10:01:47
 Job time : 2363.75 secs


```

Db 42072 CGCTTCCACACAGTCGGAGCTCAAGGACTTCGG 42106
|||||
RESULT 12
AX323453          AX323453          759 bp      DNA      linear      PAT 07-JAN-2002
LOCUS
DEFINITION      Sequence 1 from Patent WO0192308.
ACCESSION      AX323453
VERSION        AX323453.1  GI:18094216
KEYWORDS
SOURCE
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS      Passty,C.J. and Gao,Y.
TITLE        Cystine-knot polypeptides: cloaked-2 molecules and uses thereof
JOURNAL      Patent: WO 0192308-A 1 06-DEC-2001;
Amgen, Inc. (US)
FEATURES
source
BASE COUNT      125 a 282 c 244 g 108 t
ORIGIN
Query Match      10.8%; Score 69; DB 6; Length 759;
Best Local Similarity 100.0%; Pred. No. 4.3e-23;
Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 475 CTGGTGGCTCTGCAAGTCAAGCGCTTCCACACAGTCGGAGCTCAAG 534
|||||
Db 517 CTGGTGGCTCTGCAAGTCAAGCGCTTCCACACAGTCGGAGCTCAAG 576
|||||
Qy 535 GACTTCGGG 543
|||||
Db 577 GACTTCGGG 585
|||||

RESULT 13
AX342535          AX342535          2271 bp      DNA      linear      PAT 12-JAN-2002
LOCUS
DEFINITION      Sequence 1 from Patent WO0198491.
ACCESSION      AX342535
VERSION        AX342535.1  GI:18151964
KEYWORDS
SOURCE
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS      Bailemans,W., Ebeling,M., Foernzler,D., Patel,N., van Hul,W. and
Vickery,B.H.
TITLE        Osteolevin gene polymorphisms
JOURNAL      Patent: WO 0198491-A 1 27-DEC-2001;
F. HOFFMANN-LA ROCHE AG (CH) ; UNIVERSITAIRE INSTELLING ANTWERPEN
(BE)
FEATURES
source
BASE COUNT      573 a 623 c 599 g 476 t
ORIGIN
Query Match      10.8%; Score 69; DB 6; Length 2271;
Best Local Similarity 100.0%; Pred. No. 3.6e-23;
Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 475 CTGGTGGCTCTGCAAGTCAAGCGCTTCCACACAGTCGGAGCTCAAG 534
|||||
Db 493 CTGGTGGCTCTGCAAGTCAAGCGCTTCCACACAGTCGGAGCTCAAG 552
|||||

Qy 535 GACTTCGGG 543
|||||
Db 553 GACTTCGGG 561
|||||

RESULT 14
AF331844          AF331844          2296 bp      mRNA      linear      PRI 06-MAR-2001
LOCUS
DEFINITION      Homo sapiens SOST (SOST) mRNA, complete cds.
ACCESSION      AF331844
VERSION        AF331844.1  GI:13236417
KEYWORDS
SOURCE
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS      Bailemans,W., Ebeling,M., Patel,N., van Hul,E., Olson,P.,
Dioszegi,M., Lacza,C., Wuyts,W., van den Ende,J., Willems,P.,
Paes-Alves,A.F., Hill,S., Bueno,M., Ramos,F.J., Tacconi,P.,
Dijkers,F.G., Stratakis,C., Lindpaintner,K., Vickery,B.,
Foernzler,D. and Van Hul,W.
TITLE        Increased bone density in sclerosteosis is due to the deficiency of
a novel secreted protein (SOST)
JOURNAL      Hum. Mol. Genet. 10 (5), 537-543 (2001)
MEDLINE
PUBMED        11181578
REFERENCE
AUTHORS      Bailemans,W., Ebeling,M., Patel,N., Vickery,B., Foernzler,D. and Van
Hul,W.
TITLE        Direct Submission
JOURNAL      Submitted (22-DEC-2000) Medical Genetics, University of Antwerp,
Universiteitssp.ein 1, Antwerp B2610, Belgium
FEATURES
source
BASE COUNT      576 a 631 c 607 g 482 t
ORIGIN
Query Match      10.8%; Score 69; DB 9; Length 2296;
Best Local Similarity 100.0%; Pred. No. 3.6e-23;
Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 475 CTGGTGGCTCTGCAAGTCAAGCGCTTCCACACAGTCGGAGCTCAAG 534
|||||
Db 518 CTGGTGGCTCTGCAAGTCAAGCGCTTCCACACAGTCGGAGCTCAAG 577
|||||
Qy 535 GACTTCGGG 543
|||||
Db 578 GACTTCGGG 586
|||||

RESULT 15
AF326739          AF326739          2323 bp      mRNA      linear      PRI 28-FEB-2001
LOCUS
DEFINITION      Homo sapiens sclerostin mRNA, complete cds.
ACCESSION      AF326739

```

Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu

Project Information
Center project name: GRII
Center clone name: CH230-137114

Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 49288 bases at least Q40
Consensus quality: 51947 bases at least Q30
Consensus quality: 54106 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 54 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 1021: contig of 1021 bp in length
* 1022 1121: gap of unknown length
* 1122 2201: contig of 1080 bp in length
* 2202 2301: gap of unknown length
* 2302 3754: contig of 1453 bp in length
* 3755 3854: gap of unknown length
* 3855 5142: contig of 1288 bp in length
* 5143 5242: gap of unknown length
* 5243 6277: contig of 1035 bp in length
* 6278 7608: contig of 1231 bp in length
* 7609 7708: gap of unknown length
* 7709 8737: contig of 1029 bp in length
* 8738 8837: gap of unknown length
* 8838 10296: contig of 1459 bp in length
* 10297 10396: gap of unknown length
* 10397 11476: contig of 1080 bp in length
* 11477 11576: gap of unknown length
* 11577 12612: contig of 1036 bp in length
* 12613 12712: gap of unknown length
* 12713 13734: contig of 1022 bp in length
* 13735 13834: gap of unknown length
* 13835 14980: contig of 1146 bp in length
* 14981 15080: gap of unknown length
* 15081 16334: contig of 1254 bp in length
* 16335 16434: gap of unknown length
* 16435 17578: contig of 1144 bp in length
* 17579 17678: gap of unknown length
* 17679 19476: contig of 1798 bp in length
* 19477 19576: gap of unknown length
* 19577 20675: contig of 1099 bp in length
* 20676 20775: gap of unknown length
* 20776 22480: contig of 1705 bp in length
* 22481 22580: gap of unknown length
* 22581 23838: contig of 1258 bp in length
* 23839 23938: gap of unknown length
* 23939 25468: contig of 1530 bp in length
* 25469 25568: gap of unknown length
* 25569 27408: contig of 1840 bp in length
* 27409 27508: gap of unknown length
* 27509 28800: contig of 1292 bp in length
* 28801 28900: gap of unknown length
* 28901 2942: contig of 1042 bp in length
* 29943 30042: gap of unknown length
* 30043 31521: contig of 1479 bp in length
* 31522 31621: gap of unknown length
* 31622 33255: contig of 1634 bp in length
* 33256 33355: gap of unknown length

* 33356 34559: contig of 1304 bp in length
* 34560 34759: gap of unknown length
* 34760 36076: contig of 1317 bp in length
* 36077 36176: gap of unknown length
* 36177 37895: contig of 1719 bp in length
* 37896 37995: gap of unknown length
* 37996 39735: contig of 1740 bp in length
* 39736 39835: gap of unknown length
* 39836 41383: contig of 1548 bp in length
* 41384 41483: gap of unknown length
* 41484 43588: contig of 2105 bp in length
* 43589 43689: gap of unknown length
* 43690 45634: contig of 1945 bp in length
* 45635 45733: gap of unknown length
* 45734 47530: contig of 1797 bp in length
* 47531 47630: gap of unknown length
* 47631 48953: contig of 1323 bp in length
* 48954 49053: gap of unknown length
* 49054 51099: contig of 2046 bp in length
* 51100 51999: gap of unknown length
* 52029: contig of 1630 bp in length
* 52830 52929: gap of unknown length
* 52930 55451: contig of 2522 bp in length
* 55452 55551: gap of unknown length
* 55552 58096: contig of 2545 bp in length
* 58097 58196: gap of unknown length
* 58197 60127: contig of 1931 bp in length
* 60128 62306: contig of 2079 bp in length
* 62307 62406: gap of unknown length
* 62407 64228: contig of 1822 bp in length
* 64229 64328: gap of unknown length
* 64329 66422: contig of 2094 bp in length
* 66423 66522: gap of unknown length
* 66523 68670: contig of 2148 bp in length
* 68671 68770: gap of unknown length
* 68771 71676: contig of 2906 bp in length
* 71677 71776: gap of unknown length
* 71777 74664: contig of 2888 bp in length
* 74665 74764: gap of unknown length
* 74765 77233: contig of 2469 bp in length
* 77234 77333: gap of unknown length
* 77334 78578: contig of 1245 bp in length
* 78579 78678: gap of unknown length
* 78679 81968: contig of 3290 bp in length
* 81969 82068: gap of unknown length
* 82069 84418: contig of 2350 bp in length
* 84419 84518: gap of unknown length
* 84519 88073: contig of 3555 bp in length
* 88074 88173: gap of unknown length
* 88174 91080: contig of 2907 bp in length
* 91081 91180: gap of unknown length
* 91181 93633: contig of 2453 bp in length
* 93634 93733: gap of unknown length
* 93734 97324: contig of 3591 bp in length
* 97325 97424: gap of unknown length
* 97425 101506: contig of 4082 bp in length
* 101507 101606: gap of unknown length
* 101607 104898: contig of 3292 bp in length.

Query Match 16.4%; Score 104; DB 2; Length 104898;
Best Local Similarity 99.4%; Pred. No. 8.5e-41;
Matches 154; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 388 TTCCGCTGCATCCCGGATCGTACCGCGCAGCGGTGCTGTCGCCCGGGGGC 447
Db 41952 TTCCGCTGCATCCCGGATCGTACCGCGCAGCGGTGCTGTCGCCCGGGGGC 42011
Qy 448 GCGGCGCGCGCTCGCGCAAGGTGCGTCTGGTGCCTCGTGAAGTGAAGCGCTCACC 507
Db 42012 GCGGCGCGCGCTCGCGCAAGGTGCGTCTGGTGCCTCGTGAAGTGAAGCGCTCACC 42071
Qy 508 CGCTTCCACACACAGTCGGAGCTCAAGACTTCGG 542

RESULT 10

AC098160/c

LOCUS

DEFINITION

AC098160

ACCESSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

AC098160 101804 bp DNA linear HTG 12-JUL-2002
Rattus norvegicus clone CH230-60M11, *** SEQUENCING IN PROGRESS
***, 53 unordered pieces.

AC098160 3 GI:21729002
HTG; HTGS PHASE1.
Norway rat.

Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus

1 (bases 1 to 101804)
Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-Osman, F.R., Allen, C.,
Albrooks, S.L., Amarantunge, H.C., Are, J.R., Ayale, M., Banks, T.,
Barbora, J., Benton, J., Bimaga, K., Blankenburg, K., Bonnin, D.,
Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P.,
Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,
Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,
Chen, G., Chen, Z., Chowdhry, I., Christopoulos, C.,
Cleveland, C.D., Cox, C., Coyne, M.D., Dathorne, S.R., David, R.,
Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,
Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,
Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,
Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,
Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,
Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,
Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,
Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,
Hernandez, O., Hodgson, A., Hognes, M., Holloway, C., Hollins, B.,
Homi, F., Howard, S., Huber, J., Hulyk, J., Jolivet, S., Joudan, S.,
Jacobson, B., Jia, Y., Johnson, R., King, L., Korvah, J., Kovar, C.,
Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,
Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Loulseg, H.,
Lozano, R.J., Lu, X., Lucier, R., Luna, R., Ma, J.,
Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,
Massey, E., McWhiney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M.,
Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,
Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,
Nguyen, N., Nickerson, E., Nwokenkwo, S., Ogih, M., Okwuonu, G.,
Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,
Peters, L., Piekens, R., Primus, E., Pu, L.L., Quiles, M., Ran, Y.,
Rives, M., Rojas, A., Rojibokan, I., Roife, M., Ruiz, S., Savery, G.,
Scherer, S., Scott, G., Shen, H., Shooshtari, N., Sisson, I.,
Sodergren, E., Sonaite, T., Sparks, A., Stanley, H., Stone, H.,
Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,
Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
Usmani, K., Vazquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q.,
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
Williams, G., Williamson, A., Wleczek, R., Wooden, S., Worley, K.,
Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
Weinstock, G. and Gibbs, R.

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished
2 (bases 1 to 101804)
Worley, K.C.
Direct Submission
Submitted (23-OCT-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 101804)
Worley, K.C.
Direct Submission
Submitted (12-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 11, 2002 this sequence version replaced gi:17972960.
----- Genome Center
Center: Baylor College of Medicine

Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GGV
Center clone name: CH230-60M11
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 42516 bases at least Q40
Consensus quality: 46513 bases at least Q30
Consensus quality: 49790 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 53 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1
* 1001: contig of 1001 bp in length
* 1002: gap of unknown length
* 1102: contig of 1036 bp in length
* 2138: gap of unknown length
* 2238: contig of 1380 bp in length
* 3618: gap of unknown length
* 3718: contig of 1210 bp in length
* 4928: gap of unknown length
* 5028: contig of 1482 bp in length
* 6510: gap of unknown length
* 8256: contig of 1647 bp in length
* 8357: gap of unknown length
* 8577: contig of 1351 bp in length
* 9708: gap of unknown length
* 11221: contig of 1414 bp in length
* 11222: gap of unknown length
* 11322: contig of 1459 bp in length
* 12781: gap of unknown length
* 12881: contig of 1029 bp in length
* 13909: gap of unknown length
* 14010: contig of 1185 bp in length
* 15194: gap of unknown length
* 15294: contig of 1164 bp in length
* 15295: gap of unknown length
* 16458: gap of unknown length
* 16559: contig of 1318 bp in length
* 17876: gap of unknown length
* 17976: contig of 1178 bp in length
* 19154: gap of unknown length
* 19254: contig of 1715 bp in length
* 20969: gap of unknown length
* 21089: contig of 1864 bp in length
* 22933: gap of unknown length
* 23033: contig of 1187 bp in length
* 24220: gap of unknown length
* 24320: contig of 1808 bp in length
* 26128: gap of unknown length
* 26228: contig of 1467 bp in length
* 27629: gap of unknown length
* 27795: contig of 1249 bp in length
* 29044: gap of unknown length
* 29144: contig of 1057 bp in length
* 30201: gap of unknown length
* 30301: contig of 1624 bp in length
* 31925: gap of unknown length
* 32025: contig of 1498 bp in length
* 33523: gap of unknown length
* 33623: contig of 1136 bp in length
* 34759: gap of unknown length
* 34859: contig of 1368 bp in length
* 36227: gap of unknown length
* 34860

Db 60 CTCACCCGCTTCACCAACGAGTGCAGCTCAAGAGCTTGCAGCTGAGACCGCTGCGCG 1

RESULT 7

LOCUS A1113131/c

DEFINITION UI-R-C2p-nr-f-06-0-UI.s1 UI-R-C2p Rattus norvegicus cDNA clone

ACCESSION A1113131

VERSION A1113131.1 GI:3513080

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

COMMENT

315 bp mRNA linear EST 11-FEB-1999
UI-R-C2p-nr-f-06-0-UI.s1 UI-R-C2p Rattus norvegicus cDNA clone
GI:3513080
Norway rat.
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 315)
Bonaldo, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)

Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565

Email: meares@blue.weeg.uiowa.edu
Oligo-dt track not found, Not 1 site shown in beginning of sequence
is likely internal to the message. cDNA library Preparation: M.
Fatima Bonaldo, Ph.D. Clone distribution: clones will be available
through Research Genetics The following repetitive elements were
found in this cDNA sequence: 97-125, >GC-rich#low_complexity

Seq primer: M13 Forward.

FEATURES

source

Location/Qualifiers

1..315

/organism="Rattus norvegicus"

/strain="Sprague-Dawley"

/db_xref="taxon:10116"

/clone="UI-R-C2p-nr-f-06-0-UI"

/clone_1lb="UI-R-C2p"

/dev_stage="adult"

/lab_host="DH10B (Life Technologies)"

/note="Vector: pT73D-Pac (Pharmacia) with a modified

polylinker, Site 1: Not I; Site 2: Eco RI. The UI-R-C2p

library is a subtracted library derived from the UI-R-C1

library, which is a subtracted library derived from the

UI-R-C0 library. The UI-R-C0 library consisted of a

mixture of individually tagged normalized libraries

constructed from rat placenta, adult lung, brain, liver,

kidney, heart, spleen, ovary, muscle, 8, 12 and 18-day

embryo. The tag is a string of 3-5 nucleotides present

between the Not I site and the oligo-dt track which allows

identification of the library of origin of a clone within

the mixture. The subtracted library (UI-R-C2p) was

constructed as follows: PCR amplified cDNA inserts from

UI-R-C1 clones from which 3' ESTs had been derived was

used as a driver in a hybridization with the UI-R-C1

library in the form of single-stranded circles. The

remaining single-stranded circles (subtracted library) was

purified by hydroxyapatite column chromatography,

converted to double-stranded circles and electroporated

into DH10B bacteria (Life Technologies) to generate the

UI-R-C2p library. This procedure has been previously

described (Bonaldo, Lennon and Soares, Genome Research 6:

791-806, 1996)"

791-806, 1996)"

791-806, 1996)"

791-806, 1996)"

791-806, 1996)"

791-806, 1996)"

791-806, 1996)"

791-806, 1996)"

FEATURES

source

Location/Qualifiers

1..291

/organism="Rattus norvegicus"

/strain="Sprague-Dawley"

/db_xref="taxon:10116"

/clone="UI-R-BJ1"

/clone_1lb="UI-R-BJ1"

/lab_host="DH10B (Life Technologies)"

/note="Vector: pT73D-Pac (Pharmacia) with a modified

polylinker, Site 1: Not I; Site 2: Eco RI. The UI-R-BJ1

library is a subtracted library derived from the following

tissues: atrium at 16.5 dpc, ventricle at 16.5 dpc, AV

canal at 16.5 dpc, atrium at 15 dpc, ventricle at 15 dpc,

canal at 16.5 dpc, atrium at 15 dpc, ventricle at 15 dpc,

canal at 16.5 dpc, atrium at 15 dpc, ventricle at 15 dpc,

canal at 16.5 dpc, atrium at 15 dpc, ventricle at 15 dpc,

canal at 16.5 dpc, atrium at 15 dpc, ventricle at 15 dpc,

canal at 16.5 dpc, atrium at 15 dpc, ventricle at 15 dpc,

canal at 16.5 dpc, atrium at 15 dpc, ventricle at 15 dpc,

canal at 16.5 dpc, atrium at 15 dpc, ventricle at 15 dpc,

canal at 16.5 dpc, atrium at 15 dpc, ventricle at 15 dpc,

canal at 16.5 dpc, atrium at 15 dpc, ventricle at 15 dpc,

canal at 16.5 dpc, atrium at 15 dpc, ventricle at 15 dpc,

canal at 16.5 dpc, atrium at 15 dpc, ventricle at 15 dpc,

Oy 247 TACACCCGCTTCTGACAGAGCGCCATGCGCAGCGCAAGCCGATCACCAGTTGGTG 306

Db 315 TACACCCGCTTCTGACAGAGCGCCATGCGCAGCGCAAGCCGATCACCAGTTGGTG 256

Oy 307 TGTCTCCGCGCAGTGGCGGCCCGCGCGCTGCTGCCCAAGCCATCGGCGCGTGAAGTGG 366

Db 255 TGCTCGGCGCAGTGGCGGCCCGCGCGCTGCTGCCCAAGCCATCGGCGCGTGAAGTGG 196

Oy 367 TGGCGCCGGAACGAGACCGGATTTCCGTCGATCCCGATGCTACCGCGCGCAGCGGTTG 426

Db 195 TGGCGCCGGAACGAGACCGGATTTCCGTCGATCCCGATGCTACCGCGCGCAGCGGTTG 136

Oy 427 CAGCTGTGTGCG 486

Db 135 CAGCTGTGTGCG 76

Oy 487 TCGAAGTCAAGCGCTTCACCGCTTCACCAACGAGTGCAGTGCAGTGCAGTGCAGTGC 546

Db 75 TCGAAGTCAAGCGCTTCACCGCTTCACCAACGAGTGCAGTGCAGTGCAGTGCAGTGC 16

Oy 547 GAGACCG 561

Db 15 GAGACCG 1

RESULT 8

LOCUS BE111224/c

DEFINITION UI-R-BJ1-auw-b-04-0-UI.s1 UI-R-BJ1 Rattus norvegicus cDNA clone

ACCESSION BE111224

VERSION BE111224.1 GI:8503329

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

COMMENT

1 (bases 1 to 291)

Bonaldo, M.F., Lennon, G. and Soares, M.B.

Normalization and subtraction: two approaches to facilitate gene

discovery

Genome Res. 6 (9), 791-806 (1996)

Contact: Soares, MB

Program for Rat Gene Discovery and Mapping

University of Iowa

451 Eckstein Medical Research Building Iowa City, IA 52242, USA

Tel: 319 335 8250

Fax: 319 335 9565

Email: meares@blue.weeg.uiowa.edu

Oligo-dt track not found, Not 1 site shown in beginning of sequence

is likely internal to the message. cDNA library Preparation: M.B.

Soares Lab Clone distribution: clones will be available through

Research Genetics (www.resgen.com) The following repetitive

elements were found in this cDNA sequence: 97-125,

>GC-rich#low_complexity

Seq primer: M13 Forward

POLYA=No.

Location/Qualifiers

1..291

/organism="Rattus norvegicus"

/strain="Sprague-Dawley"

/db_xref="taxon:10116"

/clone="UI-R-BJ1"

/clone_1lb="UI-R-BJ1"

/lab_host="DH10B (Life Technologies)"

/note="Vector: pT73D-Pac (Pharmacia) with a modified

polylinker, Site 1: Not I; Site 2: Eco RI. The UI-R-BJ1

library is a subtracted library derived from the following

tissues: atrium at 16.5 dpc, ventricle at 16.5 dpc, AV

canal at 16.5 dpc, atrium at 15 dpc, ventricle at 15 dpc,

canal at 16.5 dpc, atrium at 15 dpc, ventricle at 15 dpc,

canal at 16.5 dpc, atrium at 15 dpc, ventricle at 15 dpc,

canal at 16.5 dpc, atrium at 15 dpc, ventricle at 15 dpc,

canal at 16.5 dpc, atrium at 15 dpc, ventricle at 15 dpc,

canal at 16.5 dpc, atrium at 15 dpc, ventricle at 15 dpc,

canal at 16.5 dpc, atrium at 15 dpc, ventricle at 15 dpc,

canal at 16.5 dpc, atrium at 15 dpc, ventricle at 15 dpc,

canal at 16.5 dpc, atrium at 15 dpc, ventricle at 15 dpc,

canal at 16.5 dpc, atrium at 15 dpc, ventricle at 15 dpc,

canal at 16.5 dpc, atrium at 15 dpc, ventricle at 15 dpc,

canal at 16.5 dpc, atrium at 15 dpc, ventricle at 15 dpc,

AV canal at 15 dpc, ventricle at 13 dpc, and adult heart.
For a detailed description of the library from which this
clone was derived, please visit our web site at
ratest.eng.uiowa.edu. The subtraction has been previously
described in (Bonaldo, Lennon and Soares, Genome Research
6:791-806, 1996)
TAG SEQ=None found"

BASE COUNT 41 a 101 c 109 g 40 t
ORIGIN

Query Match 43.0%; Score 273.4; DB 10; Length 291;
Best Local Similarity 96.2%; Pred. No. 4e-49;
Matches 280; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 271 CCATGCCGACGCGCCAGCGGTCACCGAGTTGCTGCTCCGCGCAGTGGCGCCGCG 330
DB 291 CCGTCCGCGAGTGCACCGCGGTACCGAGTTGCTGCTCCGCGCAGTGGCGCCGCG 232
QY 331 CGGCTGCTGCCCAACGCGCATCGGCGCGTGAAGTGGCGCCGCGAAGCGAGTTTC 390
DB 231 CGGCTGCTGCCCAACGCGCATCGGCGCGTGAAGTGGCGCCGCGAAGCGAGTTTC 172
QY 391 CGCTGCATCCCGGATCGCTACCGCGCGAGGGGTGACAGCTGCTGTCCCGCGCGCG 450
DB 171 CGCTGCATCCCGGATCGCTACCGCGCGAGGGGTGACAGCTGCTGTCCCGCGCGCG 112
QY 451 GCGCGCGCTGCGCGCAAGGTGCTGCTGAGGCTGTCAGAGTGCAGAGCGCTCACCCGC 510
DB 111 GCGCGCGCTGCGCGCAAGGTGCTGCTGAGGCTGTCAGAGTGCAGAGCGCTCACCCGC 52
QY 511 TTCACAAACGAGTGGAGCTCAAGAGCTTCGGGCGGAGACCGCGCGCGCG 561
DB 51 TTCACAAACGAGTGGAGCTCAAGAGCTTCGGGCGGAGACCGCGCGCGCG 1

RESULT 9
AQ171546/c 535 bp DNA linear GSS 17-OCT-1998
LOCUS HS 3088 B1 H06 MF CIT Approved Human Genomic Sperm Library D Homo
DEFINITION sapiens genomic clone Plate=3088 Col=11 Row=F, DNA sequence.

ACCESSION AQ171546
VERSION AQ171546.1 GI:3568913
KEYWORDS GSS.

SOURCE human.
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 535)

AUTHORS Mahatras,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
Keller,A., Shaker,R., Furlong,T., Young,T., Zhao,S., Adams,M.D. and
Hood,L.

TITLE Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
MEDLINE 99380589
COMMENT Contact: Mahatras GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Sequence Tagged Connector
Plate: 3088 row: P column: 11
Class: BAC ends
High quality sequence stop: 535.

FEATURES
source

1. .535
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="Plate=3088 Col=11 Row=P"
/clone_lib="CIT Approved Human Genomic Sperm Library D"
/sex="male"

/note="Organ: sperm; Vector: pBeloBAC11, BAC Clones in
E-Coli DH10B"
BASE COUNT 98 a 164 c 139 g 132 t 2 others
ORIGIN

Query Match 21.8%; Score 138.8; DB 17; Length 535;
Best Local Similarity 80.5%; Pred. No. 4.2e-20;
Matches 177; Conservative 0; Mismatches 37; Indels 6; Gaps 1;

QY 1 ATGCAGCCCTCATAGCCCGCTCATCTGCTCTACTTGTGCAAGCTGCTTGTGCT 60
DB 268 ATGCAGCTCCCACTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTA 209
QY 61 GTGAGGGCGCAGGGGTGGCAAGCTTCAGAGATGATGCAAGAGTCAATCCAGGCTT 120
DB 208 GTGAGGGCGCAGGGGTGGCAAGCTTCAGAGATGATGCAAGAGTCAATCCAGGCTC 149
QY 121 GGAGAGTACCCCGAGCTCTTCC-----TGAGAACCAACGACCATGAAACCGGCGGAG 174
DB 148 GGAGAGTACCCCGAGCTCTTCC-----TGAGAACCAACGACCATGAAACCGGCGGAG 89
QY 175 AATGAGGACAGCTTCCCATCATCTTATGACCCCAAG 214
DB 88 AAGGAGGCGGCGCTCTCACCACACTTTGAGACCAAG 49

RESULT 10
BF523030 254 bp mRNA linear EST 11-DEC-2000
LOCUS BF523030
DEFINITION UI-R-C2p-rk-a-03-0-UI-r1 UI-R-C2p Rattus norvegicus cDNA clone
UI-R-C2p-rk-a-03-0-UI 5', mRNA sequence.
ACCESSION BF523030
VERSION BF523030.1 GI:11631045
KEYWORDS EST.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE 1 (bases 1 to 254)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
COMMENT Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: meoares@blue.weeg.uiowa.edu

CDNA Library Preparation: M.B. Soares Lab Clone distribution:
clones will be available through Research Genetics (www.resgen.com)
This clone is also available through the I.M.A.G.E. Consortium at
LNNL (info@image.lnl.gov). IMAGE ID= 1793134
Seq primer: M13 Forward.

FEATURES
source

1. .254
Location/Qualifiers
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-C2p-rk-a-03-0-UI"
/clone_lib="UI-R-C2p"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies)"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-C2p
library is a subtracted library derived from the UI-R-C1
library, which is a subtracted library derived from the
UI-R-C0 library. The UI-R-C0 library consisted of a
mixture of individually tagged normalized libraries
constructed from rat placenta, adult lung, brain, liver,


```

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="FANTOM DB:0610009G22"
/db_xref="MGI:1896762"
/db_xref="taxon:10090"
/clone="0610009G22"
/sex="male"
/tissue_type="kidney"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/def_stage="adult"
70..690
/note="data source:SPR, source key:091303, evidence:ISS
homolog to HYPOHETICAL 15.3 KDA PROTEIN (FRAGMENT)
putative"
/codon_start=1
/protein_id="BAB22068.1"
/db_xref="GI:12832348"
/db_xref="MGI:1913292"
/translation="MLPPIALHSLIPLCILMKNCLAFKNDATLILYSHVKEVPAP
SSNSTLQANNGRHSSTGLDNRSVQVCELRSTKYISDQCTSPRLKELVCA
ECLPLPLPMIGGIGYGTCKWSSRSOEMRCVDRTRIDQCDDGSTRYKILTV
TACKCKRYTRONHSSHNFSVSPAPQAHRRKRAKSSKHSLS"
polya_signal
1593..1598
/note="putative"
polya_site
1612
/note="putative"
BASE COUNT      435 a      397 c      352 g      428 t
ORIGIN
Query Match      13.1%      Score 83.2;      DB 11;      Length 1612;
Best Local Similarity 52.8%      Pired. No. 4.5e-08;
Matches 284;      Conservative 0;      Mismatches 233;      Indels 21;      Gaps 4;
QY      82      GCCTTCAAGTAATGACACAGAGTGCATCCAGGCTTGAGAGTACC---CGAGCCT 138
Db      136      GCTTTAAATATGTCACAGAAATCTTTATTCATGTGTAAACCTGTCGGCA 195
QY      139      CCTCTGAGAAACAACAGCCATGAAACCGGCGGAGAGGAGAGAGCCTCC---CAC 195
Db      196      CACCCAGAGCAACAGACCCCTGATCAAGCAAGAGAGAGAGAGAGAGTTCAGTAGC 255
QY      196      CATCCCTATGAGCCCAAGATGTGTCCGAGTACAGTCCCGGAGCTGCATACCCGC 255
Db      256      ACTGAGCTGATCGAAACAGTCAGTTCAGTGGCTGCAGAGGAACTGCGGTCCACCAA 315
QY      256      TTCCTGACAGCGGCGCATGCGGCGGAGGCGGAGGCGGAGGCTGCTGCTCCGC 315
Db      316      TACATTTGAGCGGCGCATGACAGATCAAGCTCTGAGAGAGCTGTGTGCGGCGGC 375
QY      316      CAGTGGCGGCGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 363
Db      376      GAGTGTGCGGCGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 435
QY      364      TGTGTGCGGCGGCGGAGAGAGAGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 423
Db      436      TGGAGCGGAGAGAGCTCTGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 495
QY      424      GTGAGCGGCTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 493
Db      496      ATCAGCTGAGCTGTGAGAGCGG---AGCAGCGGCGACCTTACAAATACCGTGTACG 552
QY      484      TCGTGAAGTGAAGCGCTTCAACCGCTTCAACAGAGTGAAGTGAAGTGAAGTGAAG 543
Db      553      GCGTGAAGTGAAGAGTACACCGTCAAGACAGAGTCAAGAGTCAAGTTCGAAAGC 612
QY      544      CCGAGAGCGGCGGCGGCGGAGAGAGTGAAGAGCGGCGGCGGCGGCGGAGGCA 601
Db      613      GTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 670

```

```

RESULT 12
AK007893      1613 bp      mRNA      linear      HTC 19-JAN-2002
LOCUS

```

```

DEFINITION
Mus musculus 10 day old male pancreas cDNA, RIKEN full-length
enriched library, clone:1810057P14:homolog to HYPOHETICAL 15.3 KDA
PROTEIN (FRAGMENT), full insert sequence.
ACCESSION
AK007893.1      GI:12841739
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
MEDLINE
10349636
PUBMED
2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
MEDLINE
11042159
PUBMED
3
Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P.,
Kono, H., Akiyama, O., Nishi, K., Kiteama, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,
Yamamoto, R., Matsunoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsumura, S., Kawai, D.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
11076861
PUBMED
4
Kawai, J., Shingawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
Aizawa, T., Hara, A., Fukunishi, Y., Kono, H., Adachi, J., Fukuda, S.,
Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamataka, I.,
Saito, T., Okazaki, Y., Gotohori, T., Bono, H., Kasukawa, T., Saito, R.,
Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
Fleischmann, W., Gaasterland, T., Giesi, C., King, B., Kochiwa, H.,
Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Peele, G., Quackenbush, J.,
Schriml, L.M., Stebbins, F., Suzuki, R., Tomita, M.,
Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,
Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bull, C.,
Fletcher, C., Fujita, M., Gariboldi, M., Guerinich, S., Hill, D.,
Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P.,
Marchionni, L., Mashima, J., Mazzarelli, J., Mombaur, P., Nordone, P.,
Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H.,
Toyo-Oka, K., Wang, K.H., Weitz, C., Whitaker, C., Wilting, L.,
Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, H., Kohetsuki, S.
and Hayashizaki, Y.
Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)
21085660
MEDLINE
11217851
PUBMED
5 (bases 1 to 1613)
Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A.,
Arikawa, T., Baldarelli, R., Bono, H., Brownstein, M., Bull, C.,
Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T.,
Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hiraoka, T., Hori, F.,
Hume, D., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kono, H., Kouda, M., Koya, S.,
Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K.,
Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Quackenbush, J.,
Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D.,

```


QY 241 CTGACCTACACCGCTTCTGACAGACGCCCATGCGCGAGCGCGCAAGCCGGTCAACCGAG 300
DB 241 CTGACCTACACCGCTTCTGACAGACGCCCATGCGCGAGCGCGCAAGCCGGTCAACCGAG 300
QY 301 TTGGTGTGTCTCCGGCCAGATGCGCGCGCGCGCGCTGCTCCCAACGCCCATCGAGCGGTG 360
DB 301 TTGGTGTGTCTCCGGCCAGATGCGCGCGCGCGCGCTGCTCCCAACGCCCATCGAGCGGTG 360
QY 361 AAGT 420
DB 361 AAGT 420
QY 421 CGGGT 480
DB 421 CGGGT 480
QY 481 GCTGT 540
DB 481 GCTGT 540
QY 541 GGGCGGAGACCG 600
DB 541 GGGCGGAGACCG 600
QY 601 AAGCCCAACCGAGCGGAGCTGTGAGAACCGCTTACTAG 636
DB 601 AAGCCCAACCGAGCGGAGCTGTGAGAACCGCTTACTAG 636

RESULT 2

US-09-867-274-1
; Sequence 1, Application US/09867274
; Patent No. US20020106850A1
; GENERAL INFORMATION:
; APPLICANT: Paszty, Christopher
; APPLICANT: Gao, Yongming
; TITLE OF INVENTION: Cysteine Knot Polypeptides: Cloaked-2 Molecules and Uses Thereof
; FILE REFERENCE: 01017/37428
; CURRENT APPLICATION NUMBER: US/09/867,274
; PRIOR FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: US 60/208,550
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: US 60/223,542
; PRIOR FILING DATE: 2000-08-04
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0
; LENGTH: 759
; SEQ ID NO: 1
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-867-274-1

Query March 78.9%; Score 501.6; DB 10; Length 759;
Best Local Similarity 87.5%; Pred. No. 4,4e-126;

Matches 562; Conservative 0; Mismatches 74; Indels 6; Gaps 1;

QY 1 ATGCAGCCCTCACTAGCCCGCTGCTCATCTGCTACTTGTGACGCTGCTTGTGCT 60
DB 37 ATGCAGCTCCCACTAGCCCGCTGCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 96
QY 61 GTGAGGGGCGAGGGGTGCGAAGCTTCAAGATATGCAAGAGGTGATCCCGAGGCTT 120
DB 97 GTGAGGGGCGAGGGGTGCGAAGCTTCAAGATATGCAAGAGGTGATCCCGAGGCTT 156
QY 121 GAGAGTATCCCGAGGCTCTCTCC-----TGAGAACCAACGACCATGAACCGGCGAG 174
DB 157 GAGAGTATCCCGAGGCTCTCTCCCGAGAGTGTGAGAACCAACGACCATGAACCGGCGAG 216
QY 175 AATGAGGCGAGACCTCCCGACCATCTTATAGCGCCAAAGATGTGTCCGAGTACAGCTGC 234
DB 217 AATGAGGCGAGACCTCCCGACCATCTTATAGCGCCAAAGATGTGTCCGAGTACAGCTGC 276

QY 235 CGGAGCTCACTACACCGCTTCTGACAGACGCCCATGCGCGAGCGCGCAAGCCGGTCA 294
DB 277 CGGAGCTCACTACACCGCTTCTGACAGACGCCCATGCGCGAGCGCGCAAGCCGGTCA 336
QY 295 ACCGAGT 354
DB 337 ACCGAGT 396
QY 355 CGCGT 414
DB 397 CGCGT 456
QY 415 GCGAGCGGAGTCACTGCTGT 474
DB 457 GCGAGCGGAGTCACTGCTGT 516
QY 475 CTGT 534
DB 517 CTGT 576
QY 535 GACTTGT 594
DB 577 GACTTGT 636
QY 595 GAGCCCAACCGAGCGGAGCTGTGAGAACCGCTTACTAG 636
DB 637 GAGCCCAACCGAGCGGAGCTGTGAGAACCGCTTACTAG 678

RESULT 3

US-09-864-761-3098/c
; Sequence 3098, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wenheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecm1ca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: US 60/234,687
 PRIOR FILING DATE: 2000-09-21
 PRIOR APPLICATION NUMBER: US 09/608,408
 PRIOR FILING DATE: 2000-06-30
 PRIOR APPLICATION NUMBER: US 09/774,203
 PRIOR FILING DATE: 2001-01-29
 NUMBER OF SEQ ID NOS: 49117
 SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
 SEQ ID NO 30988
 LENGTH: 422
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 OTHER INFORMATION: MAP TO AC003098.1
 OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.62
 OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.6
 OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.49
 OTHER INFORMATION: EST HUMAN HIT: BE613498.1, EVALU8 9.90e-02
 OTHER INFORMATION: SWISSPROT HIT: P45646, EVALU8 4.90e-01
 OTHER INFORMATION: NT HIT: AF074705.1, EVALU8 1.00e+00
 US-09-864-761-30988

Query Match 56.3%; Score 358; DB 10; Length 422;
 Best Local Similarity 90.5%; Pred. No. 1.8e-87;
 Matches 382; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

215 ATGTGTCAGTACAGTCCGAGAGTGCATACACCCGCTTCTGACAGAGCCCAT 274
 422 AGCTGTCCAGTACAGTCCGAGAGTGCATACACCCGCTTCTGACAGAGCCCAT 363
 275 GCGCAGAGCCGAGAGTCCGAGAGTGCATACACCCGCTTCTGACAGAGCCCAT 334
 362 GCGCAGAGCCGAGAGTCCGAGAGTGCATACACCCGCTTCTGACAGAGCCCAT 303
 335 TGCTGCCCAAGCAGTCCGAGAGTGCATACACCCGCTTCTGACAGAGCCCAT 394
 302 TGCTGCCCAAGCAGTCCGAGAGTGCATACACCCGCTTCTGACAGAGCCCAT 243
 395 GCATCCCGAGTCCGAGAGTGCATACACCCGCTTCTGACAGAGCCCAT 454
 242 GCATCCCGAGTCCGAGAGTGCATACACCCGCTTCTGACAGAGCCCAT 183
 455 GCGCAGAGCCGAGAGTCCGAGAGTGCATACACCCGCTTCTGACAGAGCCCAT 514
 182 GCGCAGAGCCGAGAGTCCGAGAGTGCATACACCCGCTTCTGACAGAGCCCAT 123
 515 ACAACGAGTCCGAGAGTCCGAGAGTGCATACACCCGCTTCTGACAGAGCCCAT 574
 122 ACAACGAGTCCGAGAGTCCGAGAGTGCATACACCCGCTTCTGACAGAGCCCAT 63
 575 AGCCGAGAGCCGAGAGTCCGAGAGTGCATACACCCGCTTCTGACAGAGCCCAT 634
 62 AGCCGAGAGCCGAGAGTCCGAGAGTGCATACACCCGCTTCTGACAGAGCCCAT 3
 635 AG 636
 2 AG 1

RESULT 4
 US-09-864-761-14440/c
 Sequence 14440, Application US/09864761
 Patent No. US20020048763A1
 GENERAL INFORMATION:
 APPLICANT: Penn, Sharton G.
 APPLICANT: Rank, David R.
 APPLICANT: Hanzel, David K.
 APPLICANT: Chen, Wensheng
 TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
 FILE REFERENCE: Aemica-X-1
 CURRENT APPLICATION NUMBER: US/09/864,761
 CURRENT FILING DATE: 2001-05-23

PRIOR APPLICATION NUMBER: US 60/180,312
 PRIOR FILING DATE: 2000-02-04
 PRIOR APPLICATION NUMBER: US 60/207,456
 PRIOR FILING DATE: 2000-05-26
 PRIOR APPLICATION NUMBER: US 09/632,366
 PRIOR FILING DATE: 2000-08-03
 PRIOR APPLICATION NUMBER: GB 24263.6
 PRIOR FILING DATE: 2000-10-04
 PRIOR APPLICATION NUMBER: US 60/236,359
 PRIOR FILING DATE: 2000-09-27
 PRIOR APPLICATION NUMBER: PCT/US01/00666
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00667
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00664
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00669
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00665
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00668
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00663
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00662
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00661
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00670
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: US 60/234,687
 PRIOR FILING DATE: 2000-09-21
 PRIOR APPLICATION NUMBER: US 09/608,408
 PRIOR FILING DATE: 2000-06-30
 PRIOR APPLICATION NUMBER: US 09/774,203
 PRIOR FILING DATE: 2001-01-29
 NUMBER OF SEQ ID NOS: 49117
 SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
 SEQ ID NO 14440
 LENGTH: 392
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 OTHER INFORMATION: MAP TO AC003098.1
 OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.62
 OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.6
 OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.49
 US-09-864-761-14440

Query Match 36.5%; Score 232.4; DB 10; Length 392;
 Best Local Similarity 90.5%; Pred. No. 1.1e-53;
 Matches 248; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

363 GTGTGGGCGCCGAGAGGAGTTCGCTGATCCCGAGTGCATCCCGGAGCG 422
 392 GTGTGGGCGAGTTCGCTGATCCCGAGTGCATCCCGGAGCG 333
 423 GGTGAGTGTGTCGCGCGGAGGAGCGCGGAGTTCGCTGATCCCGGAGCG 482
 333 GGTGAGTGTGTCGCGCGGAGGAGCGCGGAGTTCGCTGATCCCGGAGCG 273
 483 CTGTGCAAGTGCAGAGTTCGCTGATCCCGGAGTTCGCTGATCCCGGAGCG 542
 272 CTGTGCAAGTGCAGAGTTCGCTGATCCCGGAGTTCGCTGATCCCGGAGCG 213
 543 GCGGAGAGCCGCGGAGGAGGAGTTCGAGAGTTCGAGAGTTCGAGAGTTCGAG 602
 212 GACCGAGAGCCGCGGAGGAGGAGTTCGAGAGTTCGAGAGTTCGAGAGTTCGAG 153
 603 AGCCAGAGGAGGAGTTCGAGAGTTCGAGAGTTCGAGAGTTCGAGAGTTCGAG 636
 152 AGCCAGAGGAGGAGTTCGAGAGTTCGAGAGTTCGAGAGTTCGAGAGTTCGAG 119

RESULT 5

US-10-152-661-430
 ; Sequence 430, Application US/10152661
 ; Publication No. US20030022835A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Watson, James D.
 ; APPLICANT: Strachan, Lorna
 ; APPLICANT: Sleeman, Matthew
 ; APPLICANT: Onrust, Rene
 ; APPLICANT: Murison, James G.
 ; APPLICANT: Kumble, Krishanand D.
 ; TITLE OF INVENTION: Compositions Isolated From Skin Cells
 ; TITLE OF INVENTION: and Methods for Their Use
 ; FILE REFERENCE: 11000.1011C5
 ; CURRENT APPLICATION NUMBER: US/10/152,661
 ; PRIOR FILING DATE: 2002-05-20
 ; PRIOR APPLICATION NUMBER: 09/866,050
 ; PRIOR FILING DATE: 2001-05-24
 ; PRIOR APPLICATION NUMBER: 60/221,232
 ; PRIOR FILING DATE: 2000-07-25
 ; PRIOR APPLICATION NUMBER: 60/206,650
 ; PRIOR FILING DATE: 2000-05-24
 ; PRIOR APPLICATION NUMBER: 09/312,283
 ; PRIOR FILING DATE: 1999-05-14
 ; PRIOR APPLICATION NUMBER: PCT/NZ99/00051
 ; PRIOR FILING DATE: 1999-04-29
 ; PRIOR APPLICATION NUMBER: 09/188,930
 ; PRIOR FILING DATE: 1998-11-09
 ; PRIOR APPLICATION NUMBER: 09/069,726
 ; PRIOR FILING DATE: 1998-04-29
 ; NUMBER OF SEQ ID NOS: 725
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 430
 ; LENGTH: 954
 ; TYPE: DNA
 ; ORGANISM: Mouse
 US-10-152-661-430

Query Match 12.8%; Score 81.6; DB 9; Length 954;
 Best Local Similarity 52.6%; Pred. No. 4e-13;

Matches 283; Conservative 0; Mismatches 214; Indels 21; Gaps 4;

QY 82 GCCTTCAGAGTATGATCCACAGAGTATCCACAGGCGCTTGAGAGTACCC---CGAGCCT 138
 DB 105 GCTTTAAATGATGACACAGAAATCTTTATTCACATGTGTTAAACCTGTCCGCA 164
 QY 139 CCTCTGAGAACAAACAGACCATGAAACCGGCGAGAGATGAGGCAACCTCCC---CAC 195
 DB 165 CACCCACAGACAGACAGACCCCTGATCAAGCCAGAGATGAGGCAAGGCAATTCAGTAGC 224
 QY 196 CATCCCTTATGACCCAAAGATGTGTCGAGTACAGCTGCCGAGCTGCACTAACCCGC 255
 DB 225 ACTGAGCTGATCGAATAAGTCAGATTCAAGTGGCTGCAAGGAACTGCCGTCACCAAA 284
 QY 256 TTCCGAGAGAGCGCCATGCGAGCGCCAGCGCGATCGGATGATGATGCTCCGGC 315
 DB 285 TACATTTGAGACGGGCAATGACACAGCATCAGCCTCTGAAGAGAGCTGTGTGCGGGC 344
 QY 316 CAGTGGCGCCCGCGCGCTGCTGCCAACGCCATTCGGGCGCG-----TGAAG 363
 DB 345 GAGTGTCTTGGCCCTGCGGTGCTTCCCAATGATCGAGAGAGGCTACGAAACAAATAC 404
 QY 364 TGTGTGGCGCCGAGACCGAGATTTCCGTGTCATCCCGATTCGCTACCGCGCGAGCGG 423
 DB 405 TGGAGCCGAGAGAGCTTCAGAGATGCGGTGTGTCAACGACCAAGCCGACCAAGAG 464
 QY 424 GTGAGCTGTGTGCGCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 483
 DB 465 ATCCAGCTGAGTGTCAAGAGAGCGC---AGCAGCGCGCATCTCAAAATACCGGTGTACG 521
 QY 484 TCGTCAAGTGAAGCGCTCAACCGCTTCCAAACACAGTGGAGCTCAAGAGACTTGGGG 543

RESULT 6

US-09-866-050A-430
 ; Sequence 430, Application US/09866050A
 ; Publication No. US2003004071A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Watson, James D.
 ; APPLICANT: Strachan, Lorna
 ; APPLICANT: Sleeman, Matthew
 ; APPLICANT: Onrust, Rene
 ; APPLICANT: Murison, James G.
 ; APPLICANT: Kumble, Krishanand D.
 ; TITLE OF INVENTION: Compositions Isolated From Skin Cells
 ; TITLE OF INVENTION: and Methods for Their Use
 ; FILE REFERENCE: 11000.1011C4U
 ; CURRENT APPLICATION NUMBER: US/09/866,050A
 ; PRIOR FILING DATE: 2001-05-24
 ; NUMBER OF SEQ ID NOS: 725
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 430
 ; LENGTH: 954
 ; TYPE: DNA
 ; ORGANISM: Mouse
 US-09-866-050A-430

Query Match 12.8%; Score 81.6; DB 9; Length 954;
 Best Local Similarity 52.6%; Pred. No. 4e-13;

Matches 283; Conservative 0; Mismatches 224; Indels 21; Gaps 4;

QY 82 GCCTTCAGAGTATGATCCACAGAGTATCCACAGGCGCTTGAGAGTACCC---CGAGCCT 138
 DB 105 GCTTTAAATGATGACACAGAAATCTTTATTCACATGTGTTAAACCTGTCCGCA 164
 QY 139 CCTCTGAGAACAAACAGACCATGAAACCGGCGAGAGATGAGGCAACCTCCC---CAC 195
 DB 165 CACCCACAGACAGACAGACCCCTGATCAAGCCAGAGATGAGGCAAGGCAATTCAGTAGC 224
 QY 196 CATCCCTTATGACCCAAAGATGTGTCGAGTACAGCTGCCGAGCTGCACTAACCCGC 255
 DB 225 ACTGAGCTGATCGAATAAGTCAGATTCAAGTGGCTGCAAGGAACTGCCGTCACCAAA 284
 QY 256 TTCCGAGAGAGCGCCATGCGAGCGCCAGCGCGATCGGATGATGATGCTCCGGC 315
 DB 285 TACATTTGAGACGGGCAATGACACAGCATCAGCCTCTGAAGAGAGCTGTGTGCGGGC 344
 QY 316 CAGTGGCGCCCGCGCGCTGCTGCCAACGCCATTCGGGCGCG-----TGAAG 363
 DB 345 GAGTGTCTTGGCCCTGCGGTGCTTCCCAATGATCGAGAGAGGCTACGAAACAAATAC 404
 QY 364 TGTGTGGCGCCGAGACCGAGATTTCCGTGTCATCCCGATTCGCTACCGCGCGAGCGG 423
 DB 405 TGGAGCCGAGAGAGCTTCAGAGATGCGGTGTGTCAACGACCAAGCCGACCAAGAG 464
 QY 424 GTGAGCTGTGTGCGCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 483
 DB 465 ATCCAGCTGAGTGTCAAGAGAGCGC---AGCAGCGCGCATCTCAAAATACCGGTGTACG 521
 QY 484 TCGTCAAGTGAAGCGCTTCAACCGCTTCCAAACACAGTGGAGCTCAAGAGACTTGGGG 543
 DB 522 GCCTGCAAGTGAAGAGGTACACCGGTCAAGCAACGATCGAGCAAACTTTGAAAGC 581
 QY 544 CCGAGACCGCGCGCGCGCGAGAGGTCGCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 601
 DB 582 GTTGTCCGACCAAGCG 639

RESULT 7

US-10-152-661-40
 ; Sequence 40, Application US/10152661
 ; Publication No. US20030022835A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Watson, James D.
 ; APPLICANT: Strachan, Lorna
 ; APPLICANT: Sleeman, Matthew
 ; APPLICANT: Onrust, Rene
 ; APPLICANT: Murison, James G.
 ; APPLICANT: Kumble, Krishanand D.
 ; TITLE OF INVENTION: Compositions Isolated From Skin Cells
 ; FILE REFERENCE: 11000.1011c5
 ; CURRENT APPLICATION NUMBER: US/10/152,661
 ; PRIOR FILING DATE: 2002-05-20
 ; PRIOR APPLICATION NUMBER: 09/866,050
 ; PRIOR FILING DATE: 2001-05-24
 ; PRIOR APPLICATION NUMBER: 60/221,232
 ; PRIOR FILING DATE: 2000-07-25
 ; PRIOR APPLICATION NUMBER: 60/206,650
 ; PRIOR FILING DATE: 2000-05-24
 ; PRIOR APPLICATION NUMBER: 09/312,283
 ; PRIOR FILING DATE: 1999-05-14
 ; PRIOR APPLICATION NUMBER: PCT/NZ99/00051
 ; PRIOR FILING DATE: 1999-04-29
 ; PRIOR APPLICATION NUMBER: 09/188,930
 ; PRIOR FILING DATE: 1998-11-09
 ; PRIOR APPLICATION NUMBER: 09/069,726
 ; PRIOR FILING DATE: 1998-04-29
 ; NUMBER OF SEQ ID NOS: 725
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 40
 ; LENGTH: 962
 ; TYPE: DNA
 ; ORGANISM: Mouse
 ; US-10-152-661-40

Query Match 12.8%; Score 81.6; DB 9; Length 962;
 Best Local Similarity 52.6%; Pred. No. 4e-13;
 Matches 283; Conservative 0; Mismatches 234; Indels 21; Gaps 4;

QY 82 GCCTTCAGGAATGATGCGACAGAGGTGATCCAGGGCTTGAGAGTACCC--CGAGCCT 138
 DB 113 GCTTTAAATAATGATGCGACAGAAATCCTTATTCATGTGTAAACCTGTCCGGCA 172
 QY 139 CTTCTGAGAACCAACAGACATGAAACCGGCGGAAATGAGAGCAGACCTCCC--CAC 195
 DB 173 CACCCAGAGCAACAGACCTGTAATCAAGCAGAAATGAGAGCGAGCATTTCACTAGC 232
 QY 196 CATCCCTATGACCCCAAGATGTGTCCAGATACAGCTGCCGAGCTGCTACACCCGC 255
 DB 233 ACTGAGCTGATGAAACAGTCAAGTTCAGAGTGGCTGCGAGGAACTGCGGTCACTAAA 292
 QY 256 TTCTTGAACAGAGCCCATGCGGAGGCGCAAGCCGCTCACAGTTGTGTCTCCGGC 315
 DB 293 TACATTTGAGACCGGACAGTGCACACAGCATGAGCCCTGTGAAGAGCTGTGTGCGGGC 352
 QY 316 CAGTGGGCGCCCGGCGGCTGTGCCCAAGCGCATGCGGGCGC-----TGAAG 363
 DB 353 GAGTGTCTGCGCCCTGCGGCTGTCCCACTGATGAGAGAGGCTACGGAACAAAGTAC 412
 QY 364 TGTGTGCGCCGAGACCGGATTTCCGCTGATCCCGGATCGTACCGCGCGAGCGG 423
 DB 413 TGAAGCGGAGAGGCTGTCAAGAGTGGCGGTGTCAACGACAAAGCCGACCCAGAGG 472
 QY 424 GTGCAAGTGTGTGCGCCGCGGCGGCGCGCGCTGCGCGCAAGGTGCGTCTGTGGCC 483
 DB 473 ATCAGCTGTGAGTGTCAAGAGCGC--AGCAGCGGCACTTACAAATTCACCGTGTACAG 529
 QY 484 TGTGCAAGTGAAGCGCTCACCGGCTTCACAAACAGTGCAGAGTCAAGAGCTTCGGG 543
 DB 530 GCTTGCAAGTGAAGAGTACACCCGTCAGACACAGAGTCCAGCCACAACTTTGAAGC 589

QY 544 CCGAGACCGCGCGCCGACAGAGGTGCGCAAGCGCGCGCCCGCGGAGAGCA 601
 DB 590 GTGTGCGCACGACGCGCCCGCCAGCACAGAGAGCGGAAGAGCCAGCAATCCA 647

RESULT 8
 US-10-152-661-209
 ; Sequence 209, Application US/10152661
 ; Publication No. US20030022835A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Watson, James D.
 ; APPLICANT: Strachan, Lorna
 ; APPLICANT: Sleeman, Matthew
 ; APPLICANT: Onrust, Rene
 ; APPLICANT: Murison, James G.
 ; APPLICANT: Kumble, Krishanand D.
 ; TITLE OF INVENTION: Compositions Isolated From Skin Cells
 ; FILE REFERENCE: 11000.1011c5
 ; CURRENT APPLICATION NUMBER: US/10/152,661
 ; PRIOR FILING DATE: 2002-05-20
 ; PRIOR APPLICATION NUMBER: 09/866,050
 ; PRIOR FILING DATE: 2001-05-24
 ; PRIOR APPLICATION NUMBER: 60/221,232
 ; PRIOR FILING DATE: 2000-07-25
 ; PRIOR APPLICATION NUMBER: 60/206,650
 ; PRIOR FILING DATE: 2000-05-24
 ; PRIOR APPLICATION NUMBER: 09/312,283
 ; PRIOR FILING DATE: 1999-05-14
 ; PRIOR APPLICATION NUMBER: PCT/NZ99/00051
 ; PRIOR FILING DATE: 1999-04-29
 ; PRIOR APPLICATION NUMBER: 09/188,930
 ; PRIOR FILING DATE: 1998-11-09
 ; PRIOR APPLICATION NUMBER: 09/069,726
 ; PRIOR FILING DATE: 1998-04-29
 ; NUMBER OF SEQ ID NOS: 725
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 209
 ; LENGTH: 962
 ; TYPE: DNA
 ; ORGANISM: Mouse
 ; US-10-152-661-209

Query Match 12.8%; Score 81.6; DB 9; Length 962;
 Best Local Similarity 52.6%; Pred. No. 4e-13;
 Matches 283; Conservative 0; Mismatches 234; Indels 21; Gaps 4;

QY 82 GCCTTCAGGAATGATGCGACAGAGGTGATCCAGGGCTTGAGAGTACCC--CGAGCCT 138
 DB 113 GCTTTAAATAATGATGCGACAGAAATCCTTATTCATGTGTGTAAACCTGTCCGGCA 172
 QY 139 CTTCTGAGAACCAACAGACATGAAACCGGCGGAAATGAGAGCAGACCTCCC--CAC 195
 DB 173 CACCCAGAGCAACAGACCTGTAATCAAGCAGAAATGAGAGCGAGCATTTCACTAGC 232
 QY 196 CATCCCTATGACCCCAAGATGTGTCCAGATACAGCTGCCGAGCTGCTACACCCGC 255
 DB 233 ACTGAGCTGATGAAACAGTCAAGTTCAGAGTGGCTGCGAGGAACTGCGGTCACTAAA 292
 QY 256 TTCTTGAACAGAGCCCATGCGGAGGCGCAAGCCGCTCACAGTTGTGTCTCCGGC 315
 DB 293 TACATTTGAGACCGGACAGTGCACACAGCATGAGCCCTGTGAAGAGCTGTGTGCGGGC 352
 QY 316 CAGTGGGCGCCCGGCGGCTGTGCCCAAGCGCATGCGGGCGC-----TGAAG 363
 DB 353 GAGTGTCTGCGCCCTGCGGCTGTCCCACTGATGAGAGAGGCTACGGAACAAAGTAC 412
 QY 364 TGTGTGCGCCGAGACCGGATTTCCGCTGATCCCGGATCGTACCGCGCGAGCGG 423
 DB 413 TGAAGCGGAGAGTGTCAAGAGTGGCGGTGTCAACGACAAAGCCGACCCAGAGG 472
 QY 424 GTGCAAGTGTGTGCGCCGCGGCGGCGCGCGCTGCGCGCAAGGTGCGTCTGTGGCC 483

APPLICANT: Haellings, Gregg A. and Adams, Mark D.
 TITLE OF INVENTION: Human CCN-Like Growth Factor
 NUMBER OF SEQUENCES: 20
 CORRESPONDENCE ADDRESS:

ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLIAN,
 CECCHI, STEWART & OLSTEIN

STREET: 6 BECKER FARM ROAD

CITY: ROSELAND

STATE: NEW JERSEY

COUNTRY: USA

ZIP: 07068

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 INCH DISKETTE

COMPUTER: IBM PS/2

OPERATING SYSTEM: MS-DOS

SOFTWARE: WORD PERFECT 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/853,625B

FILING DATE: 14-May-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/053,587

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: MULLINS, J.G.

REGISTRATION NUMBER: 33,073

REFERENCE/DOCKET NUMBER: 325800-442

TELECOMMUNICATION INFORMATION:

TELEPHONE: 201-994-1700

TELEFAX: 201-994-1744

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 900 BASE PAIRS

TYPE: NUCLEIC ACID

STRANDEDNESS: SINGLE

TOPOLOGY: LINEAR

MOLECULE TYPE: CDNA

SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-09-853-625B-1

Query Match 10.6%; Score 67.2; DB 10; Length 900;

Best Local Similarity 50.9%; Pred. No. 2.9e-09;

Matches 274; Conservative 0; Mismatches 243; Indels 21; Gaps 4;

82 GCCTTCAGAGATGATGCAAGAGTATCCAGGGCTTGAGAGTACCCCGAGC---CT 138
 184 GCTTTAAATATGATGCAAGAAATCTTTATTCATGTGTAAACCTGTTCCAGCA 243
 139 CTTCTGAGAACCAACGACCATGAAACCGGCGGAGAGTGAAGCAGACCTCC---CAC 195
 244 CACCCGAGCAACAGCAGTGAATCAAGCCAGAAATGAGGCGAGCATTTCACTAAC 303
 196 CATCCCTATAGCGCAAGATGTGTCCGAGTACAGCTGCGCGAGCTGCATACCCCGC 255
 304 ACTGAGCTGATGGAACACTCGGTTCAAGTGGTTGCGGAAACTGCTTCACCAAA 363
 256 TTCTGACAGACGCGCATGCGGAGCGGCAAGCGGTCACCGAGTTGATGTCTCGCG 315
 364 TACATCTCTATGCGCAGTGCACAGCATAGCCCTTGAAAGAGCTGTGTGTCTGCGC 423
 316 CAGTGGCGCCCGCGGCTGTGCGCAACGCCATCGGGCGGT-----GAAAG 363
 424 GAGTGTGTCGCCCTGCGCAGTGTCTCCCTAATGATGAGAGGCTATGAAACAAATGAC 483
 364 TGTGTGGCGCCGGAACGGAATTCGGTGTGATCCCGATGGCTACCGCGCGAGCGG 423
 484 TGGAGCAGAGAGAGCTCCAGAGTGGCGGTGTCTAATACAAACCTGTATCCAGAGA 543
 424 GTGAGCTGTGTGTCGCCGCGGCGCGCGCGCTCGCGCAAGGTGCTGTGTGCGCC 483
 544 ATCCAGCTGCAAGTGCAGATGGCAGC---ACAGCAGCTACAAATACAGATGATCAT 600
 484 TCGTCAAGTGCAGAGCGCTCACCCTTCACCAACAGTGGAGCTCAAGAGACTTGGGG 543

601 GCGTCAAGTGCAGAGTACACCCGCGAGCACAACGAGTTCACAACTTTGAGAGC 660
 544 CCGAGAGCCGCGCGCGCGAGAGAGGTGCGCAAGCCGCGCGCGCGCGGAGACCA 601
 661 ATGTACCTGCAAGCAGTCCAGATCACAGAGCGGAAAAAGAGCAGCAATCCA 718

RESULT 12

US-09-745-288-92/c

Sequence 92, Application US/09745288

Patent No. US20010018058A1

GENERAL INFORMATION:

APPLICANT: Reed, Steven G.

APPLICANT: Xu, Jiangchun

APPLICANT: Dillon, David C.

TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND

FILE REFERENCE: 210121.446D1

CURRENT APPLICATION NUMBER: US/09/745,288

CURRENT FILING DATE: 2000-12-19

NUMBER OF SEQ ID NOS: 101

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 92

LENGTH: 1692

TYPE: DNA

ORGANISM: Homo sapien

US-09-745-288-92

Query Match 10.3%; Score 65.6; DB 10; Length 1692;

Best Local Similarity 50.7%; Pred. No. 8e-09;

Matches 273; Conservative 0; Mismatches 244; Indels 21; Gaps 4;

82 GCCTTCAGAGATGATGCAAGAGTATCCAGGGCTTGAGAGTACCCCGAGC---CT 138
 1574 GCTTTAAATATGATGCAAGAAATCTTTATTCATGTGTAAACCTGTTCCAGCA 1515
 139 CTTCTGAGAACCAACGACCATGAAACCGGCGGAGAGTGAAGCAGACCTCC---CAC 195
 1514 CACCCGAGCAACAGCAGTGAATCAAGCCAGAAATGAGGCGAGCATTTCACTAAC 1455
 196 CATCCCTATAGCGCAAGATGTGTCCGAGTACAGCTGCGCGAGCTGCATACCCCGC 255
 1454 ACTGAGCTGATGGAACACTCGGTTCAAGTGGTTGCGGAACTGCTTCACCAAA 1395
 256 TTCTGACAGACGCGCATGCGGAGCGGCAAGCCGCTACCGAGTTGATGTCTCGCG 315
 1394 TACATCTCTATGAGCGCAGTGCACAGCATACGCCCTTGAAAGAGCTGTGTGTCTGCGC 1335
 316 CAGTGGCGCCCGCGGCTGTGCGCAACGCCATCGGGCGGT-----GAAAG 363
 1334 GAGTGTGTCGCCCTGCGCAGTGTCTCCCTAATGATGAGAGGCTATGAAACAAATGAC 1275
 364 TGTGTGGCGCCGGAACGGAATTCGGTGTGATCCCGATGGCTACCGCGCGAGCGG 423
 1274 TGGAGCAGAGAGAGCTCCAGAGTGGCGGTGTCTAATACAAACCTGTATCCAGAGA 1215
 424 GTGAGCTGTGTGTCGCCGCGGCGCGCGCGCTCGCGCAAGGTGCTGTGTGCGCC 483
 1214 ATCCAGCTGCAAGTGCAGATGGCAGC---ACAGCAGCTACAAATACAGATGATCAT 1158
 484 TCGTCAAGTGCAGAGCGCTCACCCTTCACCAACAGTGCAGAGCTCAAGAGACTTGGGG 543
 1157 GCGTCAAGTGCAGAGAGTGCACCGCGGAGCAGACAGATGCAAGTCAAGTCAAGAGC 1098
 544 CCGAGAGCCGCGCGCGCGCAGAGAGGTGCGCAAGCCGCGCGCGCGCGGAGACCA 601
 1097 ATGTACCTGCAAGCAGTCCAGATCACAGAGCGGAAAAAGAGCAGCAATCCA 1040

RESULT 13

US-09-803-719-2375

Sequence 2375, Application US/09803719

```

; Publication No. US20030044783A1
; GENERAL INFORMATION:
; APPLICANT: Williams, Lewis T.
; APPLICANT: Escobedo, Jaime
; APPLICANT: Imnis, Michael A.
; APPLICANT: Garcia, Pablo Dominguez
; APPLICANT: Sudduth-Klinger, Julie
; APPLICANT: Reinhard, Christoph
; APPLICANT: Giese, Klaus
; APPLICANT: Randazzo, Filippo
; APPLICANT: Kennedy, Giulia C.
; APPLICANT: Pot, David
; APPLICANT: Kassam, Altaf
; APPLICANT: Lamson, George
; APPLICANT: Drmanac, Radoje
; APPLICANT: Crkvenjakov, Radomir
; APPLICANT: Dickson, Mark
; APPLICANT: Drmanac, Snezana
; APPLICANT: Labat, Ivan
; APPLICANT: Leehkowitz, Dena
; APPLICANT: Kita, David
; APPLICANT: Garcia, Veronica
; APPLICANT: Jones, Lee William
; APPLICANT: Stache-Crain, Birgit
; TITLE OF INVENTION: Human Genes and Gene Products
; FILE REFERENCE: 1624.002
; CURRENT APPLICATION NUMBER: US/09/803,719
; CURRENT FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 60/188,609
; PRIOR FILING DATE: 2000-03-09
; NUMBER OF SEQ ID NOS: 2396
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2375
; LENGTH: 373
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-803-719-2375

Query Match          6.9%; Score 44.2; DB 9; Length 373;
Best Local Similarity 53.8%; Pred. No. 0.0045;
Matches 91; Conservative 0; Mismatches 78; Indels 0; Gaps 0.

QY 396 CATCCCGGATCGCTACCGGCGCAGCGGGGTGACGCTGTGTGCCCGGGGGGGCGGGCGCC 455
Db 93 CAAGCAGGAGCAGCTACCGCTGTCAGGGACTGCAAGGCTCTGGGGGGCGGGGGGCT 152
QY 456 GCGCTCGCCCAAGTGCCTGCTGTGGCTCTGTGCAAGTCAAGCGGCTCAACCGCTTCCA 515
Db 153 GCGCGAGCCCCCAGCCCCGCTTTGTGCCCTCTCAGACTCAAGTCAAGCTTCCCCCTTCA 212
QY 516 CAACCAGTGAGGCTCAAGGACTTCGGGCGCGAGACCGCGCGGCGCGAG 564
Db 213 GAGCAACCCCGACACCGGGAAGTCTTGGCCAGGGAGTGGGGAGCGGGGAG 261

RESULT 14
US-09-803-719-2377
; Sequence 2377, Application US/09803719
; Publication No. US20030044783A1
; GENERAL INFORMATION:
; APPLICANT: Williams, Lewis T.
; APPLICANT: Escobedo, Jaime
; APPLICANT: Imnis, Michael A.
; APPLICANT: Garcia, Pablo Dominguez
; APPLICANT: Sudduth-Klinger, Julie
; APPLICANT: Reinhard, Christoph
; APPLICANT: Giese, Klaus
; APPLICANT: Randazzo, Filippo
; APPLICANT: Kennedy, Giulia C.
; APPLICANT: Pot, David
; APPLICANT: Kassam, Altaf
; APPLICANT: Lamson, George
; APPLICANT: Drmanac, Radoje

```

```

; APPLICANT: Ctkvenjakov, Radomir
; APPLICANT: Dickson, Mark
; APPLICANT: Drmanac, Snezana
; APPLICANT: Labat, Ivan
; APPLICANT: Leshkowitz, Dena
; APPLICANT: Kita, David
; APPLICANT: Garcia, Veronica
; APPLICANT: Jones, Lee William
; APPLICANT: Strache-Grain, Birgit
; TITLE OF INVENTION: Human Genes and Gene Products
; FILE REFERENCE: 1624.002
; CURRENT APPLICATION NUMBER: US/09/803.719
; CURRENT FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 60/188,609
; PRIOR FILING DATE: 2000-03-09
; NUMBER OF SEQ ID NOS: 2396
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2377
; LENGTH: 426
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-803-719-2377

Query Match
Best Local Similarity 53.8%; Score 44.2; DB 9; Length 426;
Matches 91; Conservative 0; Mismatches 78; Indels 0; Gaps 0;

336 CATCCCGATGCTACCGCGCGCAGCGGCTGAGTGTCTGCCCCGGGGGCGCGGCC 455
Db 54 CAAGCAGAGAGAGCTACCGCTGCGCAGGGACTGACGGGCTCTGGGGGCGCGGGGCT 113
QY 456 GGGCTCGGCGCAAGGCGCTGTGTTGGCTCGGCGCAAGTGCAGCGCTCACCGCTTCCA 515
Db 114 GGGCAGGCGCCCGAGCGCCCTTTGTGCCCCCTCAGACTGCAGACTGACCTTCCCTTCAA 173
QY 516 CAACCATGTGAGCTCAAGAGACTTTCGGGCGGAGACCGCGCGCCGAG 564
Db 174 GAGCAACCCCGACGACCGGGAGTCTTGGCCAGGGATGGGAGCGGGGAG 222

RESULT 15
US-09-822-830A-91
; Sequence 91, Application US/09822830A
; Patent No. US20020142952A1
; GENERAL INFORMATION:
; APPLICANT: Genetec Institute, Inc.
; APPLICANT: Wong, Gordon G.
; APPLICANT: Clark, Hilary
; APPLICANT: Fechtel, Kim
; APPLICANT: Agostino, Michael J.
; APPLICANT: Howes, Steven H.
; APPLICANT: Resnick, Richard J.
; APPLICANT: Gulukota, Kamalakara
; APPLICANT: Graham, James R.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS
; FILE REFERENCE: GIN 6402
; CURRENT APPLICATION NUMBER: US/09/822.830A
; CURRENT FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/195,604
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 631
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 91
; LENGTH: 1951
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-822-830A-91

Query Match
Best Local Similarity 53.8%; Score 44.2; DB 10; Length 1951;
Matches 91; Conservative 0; Mismatches 78; Indels 0; Gaps 0;

336 CATCCCGATGCTACCGCGCGCAGCGGCTGAGTGTCTGCCCCGGGGGCGCGGCC 455

```

```

Db      918  CAAACAGAGAGAGCTACGGCTGCAAGGACTGCAAGGCTTGAGGGCCCGGGCGGGGGCT 977
Qy      456  GCGCTCGGCAAGGTGGCTGTGAGCTGTGCAAGTGCAGAGGCTCAACCGCTTCCA 515
Db      978  GCGCAGGCCCCCAGCCGCTTTGTGCCCCCTCAGACTGCAAGCTACGCTTCCCTTCAA 1037
Qy      516  CAAACAGTGGAGCTCAAGACTTGGGGCCGAGACCGCGCGCGCAG 564
Db      1038  GAGCAACCCCAAGCAGCGGAGTCTTGGCCAGGATGGGAGCGGGGAG 1086

```

Search completed: March 29, 2003, 01:33:24
 Job time : 79.1376 secs

PF 29-MAY-2001; 2001WO-US17478.
 XX
 PR 01-JUN-2000; 2000US-20850P.
 PR 04-AUG-2000; 2000US-223542P.
 XX
 PA (AMGE-) AMGEN INC.
 PI
 PI Paszy CJ, Gao Y;
 XX
 XX WPI: 2002-114325/15.
 DR P-PSDB; ABB07208, ABB07210.
 XX
 PT New human and mouse cysteine-knot polypeptide designated as Cloaked-2,
 PT for treating or preventing kidney, heart (e.g. myocardial infarction)
 PT or liver (e.g. hepatitis) diseases -
 XX
 PS Claim 1; Fig 2; 170pp; English.
 XX
 CC The invention relates to polypeptides comprising a cysteine knot motif
 CC and designated as Cloaked-2, derived from human and mouse. The cloaked-2
 CC polypeptides can be expressed by standard recombinant methodology. The
 CC cloaked-2 polynucleotides are useful in gene therapy and antisense
 CC therapy. The cloaked-2 polypeptides and polynucleotides are useful for
 CC treating, preventing, ameliorating or detecting diseases and disorders of
 CC the kidney (e.g. anemia, hypertension or low blood pressure), heart (e.g.
 CC cardiac hypertrophy, congestive heart failure, myocardial infarction,
 CC arrhythmias, atherosclerosis, hypertension or low blood pressure),
 CC skeletal muscle (e.g. muscular dystrophy or cachexia), placenta (e.g.
 CC congenital abnormalities or miscarriage), liver (e.g. hepatitis or
 CC cirrhosis), pancreas (e.g. diabetes or pancreatitis), thyroid (e.g.
 CC Grave's disease or myxedema) or adrenal cortex (e.g. Cushing's disease
 CC or Addison's disease), homeostasis or metabolic diseases (e.g. obesity,
 CC cancer or myopathies), infections, or autoimmune diseases. Selective
 CC binding agents may be used to modulate the biological activities of
 CC cloaked-2 polypeptides or to detect Cloaked-2 polypeptide levels in a
 CC sample. Transgenic non-human animals are useful for drug candidate
 CC screening. The present sequence represents a cDNA encoding the mouse
 CC cloaked-2 polypeptide.
 CC
 XX
 SO Sequence 636 BP; 114 A; 224 C; 207 G; 91 T; 0 other;
 Query Match 100.0%; Score 636; DB 24; Length 636;
 Best Local Similarity 100.0%; Pred. No. 8.1e-124;
 Matches 636; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATGAGCCCTCCTAGAGCCCGGCTGATCTGCTAGTGGACGGCTTCTGTGCT 60
 DB 1 ATGAGCCCTCCTAGAGCCCGGCTGATCTGCTAGTGGACGGCTTCTGTGCT 60
 QY 61 GTGAGGGCCAGGGGGTGGCAAGCTTCAAGATGATGCACAGAGTCAATCCAGGCTT 120
 DB 61 GTGAGGGCCAGGGGGTGGCAAGCTTCAAGATGATGCACAGAGTCAATCCAGGCTT 120
 QY 121 GGAAGTACCCCGAGCTCTCTGAGAACCAACCAACATGAACCGGGCGAGATGGA 180
 DB 121 GGAAGTACCCCGAGCTCTCTGAGAACCAACCAACATGAACCGGGCGAGATGGA 180
 QY 181 GGCAGACCTCCCAACATCCCTATGAGCGCCAAAGATGTCCTCGGTACAGTCCGGGAG 240
 DB 181 GGCAGACCTCCCAACATCCCTATGAGCGCCAAAGATGTCCTCGGTACAGTCCGGGAG 240
 QY 241 CTGCACTACACCCGCTTCTGAGACAGAGCCCATGCGGCGCAAGCCGGTCAACCGAG 300
 DB 241 CTGCACTACACCCGCTTCTGAGACAGAGCCCATGCGGCGCAAGCCGGTCAACCGAG 300
 QY 301 TTGCTGTCTCCGCGACAGTGGCGCGCGCGCTGCTGCTCCCAACGCCATCGCGCGCTG 360
 DB 301 TTGCTGTCTCCGCGACAGTGGCGCGCGCGCGCTGCTGCTCCCAACGCCATCGCGCGCTG 360
 QY 361 AAGGTGGGGCCCGGAACGAGCCGATTTCCGCTGACATCCCGATGCTACCGCGCGAG 420
 DB 361 AAGGTGGGGCCCGGAACGAGCCGATTTCCGCTGACATCCCGATGCTACCGCGCGAG 420

QY 421 CGAGTGCAGACTCTGTGCCCGGGGGCGGCGCGCGCTGCGCCAAAGTGCCTGTGTG 480
 DB 421 CGAGTGCAGACTCTGTGCCCGGGGGCGGCGCGCGCTGCGCCAAAGTGCCTGTGTG 480
 QY 481 GCCTGTGTGAATGTGAACGCTTCAACCGCTTCCACAAACGATGGAGCTCAAGACTTC 540
 DB 481 GCCTGTGTGAATGTGAACGCTTCAACCGCTTCCACAAACGATGGAGCTCAAGACTTC 540
 QY 541 GGGCCGAGAGACCGCGCGCGCGAGAAAGGATGCAAGCGCGCGCGCGCGCGAGAGCC 600
 DB 541 GGGCCGAGAGACCGCGCGCGCGAGAAAGGATGCAAGCGCGCGCGCGCGAGAGCC 600
 QY 601 AAAGCCAAACAGCGAGCTGAGAACCGCTACTAG 636
 DB 601 AAAGCCAAACAGCGAGCTGAGAACCGCTACTAG 636
 RESULT 2
 AAA29058
 ID AAA29058 standard; cDNA; 638 BP.
 XX
 AC AAA29058;
 XX
 DT 12-SEP-2000 (first entry)
 XX
 DE Murine TGF-beta binding protein (BBER) cDNA.
 KW osteoparh; transforming growth factor-beta; TGF-beta; binding protein;
 KW BBER; gene therapy; antisense therapy; fracture; bone mineralization; ss.
 XX
 OS Mus musculus.
 XX
 FH Key location/Qualifiers
 FT CDS 1..636
 FT /tag= a
 FT /product= TGF-beta_binding_protein
 XX
 PN WO200032773-A1.
 PD 08-JUN-2000.
 XX
 PF 24-NOV-1999; 99WO-US27990.
 XX
 PR 27-NOV-1998; 98US-0110283.
 XX
 PA (DARW-) DARWIN DISCOVERY LTD.
 XX
 PI Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepker BW;
 PI Van Ness J, Winkler Dg;
 XX
 DR WPI: 2000-412321/35.
 PT P-PSDB; AAY96432.
 PT
 PT Nucleic acids (1) encoding a transforming growth factor beta binding
 PT protein, useful for identifying agents for treating osteopenia,
 PT osteoporosis and fractures
 XX
 PS Claim 1; Page 123; 162pp; English.
 XX
 CC This cDNA encodes a murine transforming growth factor-beta (TGF-beta)
 CC binding protein designated mBER. The cDNA and protein may be used for
 CC prevention, treatment and diagnosis of diseases associated with
 CC inappropriate BBER expression. For example, they may be used to treat
 CC disorders associated with decreased TGF-beta BP expression. The cDNA or
 CC vectors may be administered to treat diseases by rectifying mutations or
 CC deletions in a patient's genome that affect the activity of BBER by
 CC expressing inactive proteins or to supplement the patient's own production
 CC of BBER polypeptides. The nucleic acids may be used for recombinant
 CC production of BBER, gene therapy, antisense therapy, as probes for
 CC diagnostic assays and for functional studies. BBER may be used to raise
 CC antibodies and for identification of BBER modulators. BBER antagonists
 CC may be used to increase bone mineral content for the treatment of
 CC disorders such as osteopenia, osteoporosis, fractures and other disorders

CC associated with low mineral content.

SQ Sequence 638 BP; 114 A; 224 C; 209 G; 91 T; 0 other;

Query Match	99.7%;	Score 634.4;	DB 21;	Length 638
Best Local Similarity	99.8%;	Pred. No. 1.8e-123;		
Matches 635; Conservative	0;	Mismatches 1;	Indels 0;	

Qy	1	ATGAGGCTCTACTAGGCGCCGTGCTCATCTGCTTACTTGGCAAGCTT3CCTTCTGAGT	60
Db	1	ATGAGGCTCTCACTAGCCCGGTGCTCATCTGCTTACTTGGCAAGCTT3CCTTCTGAGT	60
Qy	61	GTGAGGAGCCAGGGGTGGCAAGCTTCAAGATGATGCCAAGAGTCAATCCAGGGCTT	120
Db	61	GTGAGGGGCGCAGGGGTGGCAAGCTTCAAGATGATGCCAAGAGTCAATCCAGGGCTT	120
Qy	121	GGAGAGTACCCCGAGGCTCTCCCTCGAACAACAACGACCAATGAAACCGG3GGGAAATGGA	180
Db	121	GGAGAGTACCCCGAGGCTCTCTCGAACAACAACGACCAATGAAACCGG3GGGAAATGGA	180
Qy	181	GGCAGAGCTCCCGACCAATCCCTTATGACGCGAAAGATGTGTCCGAGTACAGCTGCGCGAG	240
Db	181	GGCAGAGCTCCCGACCAATCCCTTATGAGCGCAAGGTGTGTTCGAGTACAGCTGCGCGAG	240
Qy	241	CTGCACTACACCCCGCTTCTTGACAGAGGCCCATGCGCGACGCGCAAGCCGGTCAACGAG	300
Db	241	CTGCACTACACCCCGCTTCTTGACAGAGGCCCAATGCGCGACGCGCAAGCCGGTCAACGAG	300
Qy	301	TTGGTGTGCTCCGCGCAAGTGCGGCGCGCGCGGTGCTGCGCAACGCGCATCCGCGGGTGG	360
Db	301	TTGGTGTGCTCCGCGCAAGTGCGGCGCGCGCGGTGCTGCGCAACGCGCATCCGCGGGTGG	360
Qy	361	AAAGTGTGGCGCCCGAAACGAGCCGGAATTTCCGCTGCATCCCGAATCGTACCGCGCGAG	420
Db	361	AAAGTGTGGCGCCCGAAACGAGCCGGAATTTCCGCTGCATCCCGAATCGTACCGCGCGAG	420
Qy	421	CGGGTGGAGCTGCTGTGCGCCCGGGGGGGCGCGCGCGCGCTCGGCAAGGTGGCTTGGTGG	480
Db	421	CGGGTGGAGCTGCTGTGCGCCCGGGGGGGCGCGCGCGCGCTCGGCAAGGTGGCTTGGTGG	480
Qy	481	GCTCTGTGCAAGTGCAAGCGCGCTTACCCGCTTCCACAAACAGTTCGAGATCAAGAGACTTC	540
Db	481	GCTCTGTGCAAGTGCAAGCGCGCTTACCCGCTTCCACAAACAGTTCGAGATCAAGAGACTTC	540
Qy	541	GGGCGCGAGACCGCGCGCGCGCAGAAAGGTGCAAGCTCGCGGCCCGGCG3CCCGGGAGGCC	600
Db	541	GGGCGCGAGACCGCGCGCGCGCAGAAAGGTGCAAGCTCGCGGCCCGGCG3CCCGGGAGGCC	600
Qy	601	AAAGCAACCAAGCGCGAGGCTGGAAAGAGGCTCAATGAG	636
Db	601	AAAGCAACCAAGCGCGAGGCTGGAAAGAGGCTCAATGAG	636

RESULT 3
AAA29059

AC	AAA29059;	
XX		
DT	12-SEP-2000	(first entry)
XX		
DE	Rat TGF-beta binding protein (BBER) cDNA.	
XX		
KW	osteopachic; transforming growth factor-beta; TGF-beta; binding protein; BBER; gene therapy; antiscense therapy; fracture; bone mineralization; ss	
XX		
OS	Rattus norvegicus.	
XX		
XX	Key	Location/Qualifiers
FT	CDS	33..674
FT		/*tag= a
XX		/product= TGF-beta_binding_protein

PN WO200032773-A1.

PD 08-JUN-2000

PF 24-NOV-1999; 99WO-US27990.

PR 27-NOV-1998; 98US-0110283.
VY

PA (DARW-) DARWIN DISCOVERY LTD.
XX
XX

PI	BRUNKOW ME,	GALAS DJ,	KOVACEVICH B,	MULLIGAN JT,	PAEPEL BW,
PI	VAN NESS J,	WINKLER DG:			

DR WPI: 2000-412321/35.

XX
XX

PT protein, useful for

Claim 1; Page 125; 162pp; English.

CC This cDNA encodes a rat transformi
CC binding protein of approx 20000

CC prevention treatment and diagnosis of diseases associated with
CC inappropriate BEEr expression. For example, they may be used to treat
CC disorders associated with decreased TGF- β BP expression. The cDNA or
CC vectors may be administered to treat diseases by rectifying mutations or
CC deletions in a patient's genome that affect the activity of BEEr by
CC expressing inactive proteins or to supplement the patients own production
CC of BEEr polypeptides. The nucleic acids may be used for recombinant
CC production of BEEr, gene therapy, antisense therapy, as probes for
CC diagnostic assays and for functional studies. BEEr may be used to raise
CC antibodies and for identification of BEEr modulators. BEEr antagonists
CC may be used to increase bone mineral content for the treatment of
CC disorders such as osteopenia, osteoporosis, fractures and other disorders
CC associated with low mineral content.

SQ Sequence 674 BP; 128 A; 240 C; 208 G; 98 T; 0 other;

Query Match	87.7%;	Score 557.6;	DB 21;	Length 674;
Best Local Similarity	93.0%;	Pred. No. 1.8e-107;		
Matches 597;	Conservative	0;	Mismatches 39;	Indels 6; Gaps 1

Qy	1	ATGAGGCGCTCACTAGAGCCGGTGGCTCATCGCCTATTGAGACGGTGGCTTCTGTC	60
Db	33	ATGAGGCTCTACTAGAGCCCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT	92
Qy	61	GTGAGGCGCCAGGGGTGGCAAGCTTTCAGGAATGATCCACAGAGGTGCATCCAGGCTT	120
Db	93	GTGAGAGCCAGGGGTGGCAAGCTTTCAGGAATGATCCACAGAAATCATCCGGGCTC	152
Qy	121	GGAGGTAACCCCGAGCTCTCT-----CTGAGAACCAACGAGCATGAACCGGGCGAG	174
Db	153	AGAGAGTAACCCAGAGCTCTCTCAGGAATGAGAACCAACGAGCATGAACCGGGCGAG	212
Qy	175	AATGAGGCGAGACTCTCCCAACATCTCCCTTAGACGCGCAAGTGTGTCCAGTACAGCTGC	234
Db	213	AACGAGGCGAGACCCCCCCCAACATCTCTTAGACACCAAAAGCTGTCCAGTACAGCTGC	272
Qy	235	CGCAGAGTCGACTACACCCGGTTCCTGACAGACGGGCCATGCCGACGGCCAAAGCGGTC	294
Db	273	CGCAGAGTCGACTACACCCGCTTGTGACCGAGCGGCCGTGCCAGTGCCAAAGCGGTC	332
Qy	295	ACCGAGTGTGGTGTCTCCGGCCAGTGGGGCCCCGGGGGTGTGCTGCCCAACGCAATCGGG	354
Db	333	ACCGAGTGTGGTGTCTCGGGCCAGTGGGGCCCCGGGGGTGTGCTGCCCAACGCAATCGGG	392
Qy	355	CGCGTGAAGTGTGGCCGCCCGGAACGGAATTTCCGCTGCATCCCGGATTCGCTACCGC	414
Db	393	CGCGTGAAGTGTGGCCGCCCGGAACGGAATTCCTCGCTGCATCCCGGATTCGCTACCGC	452
Qy	415	CGCGAGCGGGTGCAGCTGTGTGCCCCGGGGGGCGCGGCCGCTGCAGCAAGTGTGCT	474

DB 453 GCGGAGCGGCTGAGCTCTGTGCTCCGCGCGCGCGCGCGCTCGCGCAAGTGCCT 512

QY 475 CTGTGGCTCTGTGCAAGTCAAGCGCTTCCAGCAACAGTGCAGTCAAG 534

DB 513 CTGTGGCTCTGTGCAAGTCAAGCGCTTCCAGCAACAGTGCAGTCAAG 572

QY 535 GACTTGGGCGGAGACCGCGCGCGCGCGCAAGAGTGCAGACCGCGCGCGCG 594

DB 573 GACTTGGGCGGAGACCGCGCGCGCGCGCAAGAGTGCAGACCGCGCGCGCG 632

QY 595 GGAGCCAAAGCCAAAGCGGAGCTGAGAAAGCTACTAG 636

DB 633 GGAGCCAAAGCCAAAGCGGAGCTGAGAAAGCTACTAG 674

RESULT 4
AAA29057
ID AAA29057 standard; cDNA; 642 BP.

AC AAA29057;
XX
DT 12-SEP-2000 (first entry)
XX
DE Vervet TGF-beta binding protein (BBER) cDNA.
XX
KM osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
KW BBER; gene therapy; antisense therapy; fracture; bone mineralization; ss.
XX
OS Cercopithecus pygerythrus.
XX
FH
FT Key Location/Qualifiers
FT CDS 1..642
FT /tag= a
FT /product= TGF-beta_binding_protein

MO200032773-A1.
XX
PD 08-JUN-2000.
XX
PF 24-NOV-1999; 99MO-US27990.
XX
PR 27-NOV-1998; 98US-0110283.
XX
PA (DARM-) DARWIN DISCOVERY LTD.
PI Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepker BW;
PI Van Ness J, Winkler DG;
XX
DR WPI: 2000-412321/35.
XX
DR P-PSDB; AAY96431.
XX
PT Nucleic acids (1) encoding a transforming growth factor beta binding
PT protein, useful for identifying agents for treating osteopenia,
PT osteoporosis and fractures
XX
PS Claim 1; Page 122; 162pp; English.

XX
CC This cDNA encodes a vervet transforming growth factor-beta (TGF-beta)
CC binding protein designated vBER. The cDNA and protein may be used for
CC prevention, treatment and diagnosis of diseases associated with
CC inappropriate BBER expression. For example, they may be used to treat
CC disorders associated with decreased TGF-beta BP expression. The cDNA or
CC vectors may be administered to treat diseases by rectifying mutations or
CC deletions in a patient's genome that affect the activity of BBER by
CC expressing inactive proteins or to supplement the patient's own production
CC of BBER polypeptides. The nucleic acids may be used for recombinant
CC production of BBER, gene therapy, antisense therapy, as probes for
CC diagnostic assays and for functional studies. BBER may be used to raise
CC antibodies and for identification of BBER modulators. BBER antagonists
CC may be used to increase bone mineral content for the treatment of
CC disorders such as osteopenia, osteoporosis, fractures and other disorders
CC associated with low mineral content.

XX
SQ Sequence 642 BP; 113 A; 228 C; 210 G; 91 T; 0 other;

Query Match 79.4%; Score 504.8; DB 21; Length 642;
Best local similarity 87.9%; Pred. No. 1,9e-96;
Matches 564; Conservativity 0; Mismatches 72; Indels 6; Gaps 1;

QY 1 ATGAGCCCTCAGTACAGCCCGTGCCTCATCTGCTTGTGACAGCTGCTTGTGCT 60

DB 1 ATGAGCTTCCACTGAGCCCTGTGTCTGTGTCTGCTGCTGCTGCTGCTGCTGCT 60

QY 61 GTGAGGCGCCAGGGGTGCAAGCCTTCAAGAAATGATGCCAAGAGTCAATCCAGGCTT 120

DB 61 GTGAGGCGCCAGGGGTGCAAGCCTTCAAGAAATGATGCCAAGAGTCAATCCAGGCTT 120

QY 121 GAGAGTACCCCGAGCTCTTC-----TGAGAAACAACAGACATGAACCGGCGGAG 174

DB 121 GGAGAGTACCCCGAGCTCTTCACCGAGCTGAGAAACAACAGACATGAACCGGCGGAG 180

QY 175 AATGAGGCGAGACCTCCCGACCATCCCTATGAGCGCAAAAGTGTGTCGAGTACAGCTGC 234

DB 181 AATGAGGCGGCGGCTCTCCACACCCCTTTGAGACCAAAAGCTGTCCGAGTACAGCTGC 240

QY 235 CGCAGCTGCACTACACCGCTTCTTGAACAAGCGCCATGCGCAGCGCAAGCCGCTC 294

DB 241 CGAGAGCTGCACTTCAACCGCTTCAAGTGAAGTGGCCCGTGCAGCGCAAGCCAGTGC 300

QY 295 ACCGAGTTGTGTGCTCCGCGCAAGTGCAGCGCGCGCGGCTGCTGCTGCTGCTGCT 354

DB 301 ACCGAGTTGTGTGCTCCGCGCAAGTGCAGCGCGCGCGGCTGCTGCTGCTGCTGCT 360

QY 355 GCGCGAAAGTGTGTGCGCGCGCGCAAGCGGATTCGCTGCTGCTGCTGCTGCTGCTGCT 414

DB 361 GCGCGAAAGTGTGTGCGCGCGCGCAAGCGGATTCGCTGCTGCTGCTGCTGCTGCTGCT 420

QY 415 GCGCAGCGGCTGCAAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 474

DB 421 GCGCAGCGGCTGCAAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480

QY 475 CTGTGGCTCTGTGCAAGTCAAGCGCTTCCAGCAACAGTGCAGTCAAG 534

DB 481 CTGTGGCTCTGTGCAAGTCAAGCGCTTCCAGCAACAGTGCAGTCAAG 540

QY 535 GACTTGGGCGGAGACCGCGCGCGCGCGCAAGAGTGCAGACCGCGCGCGCGCGCG 594

DB 541 GACTTGGGCGGAGACCGCGCGCGCGCGCAAGAGTGCAGACCGCGCGCGCGCGCG 600

QY 595 GGAGCCAAAGCCAAAGCGGAGCTGAGAAAGCTACTAG 636

DB 601 GGAGCCAAAGCCAAAGCGGAGCTGAGAAAGCTACTAG 642

RESULT 5
AAA94051
ID AAA94051 standard; DNA; 642 BP.

AC AAA94051;
XX
DT 15-JAN-2001 (first entry)
XX
DE Human DAN/Cerberus-related protein 6 (hDCR6) cDNA exons 1 and 4.
XX
KW Human; DNA/Cerberus-related protein 6; hDCR6; morphogenic protein;
KW antagonist; BMP; cell growth; cell differentiation; bone formation;
KW gene therapy; ss.
XX
OS Homo sapiens.
XX
FH
FT Key Location/Qualifiers
FT CDS 1..642
FT /tag= a
FT /product= "hDCR6 #2"
FT /partial

XX	PN	MOZ00055193-A2.
XX	PD	21-SEP-2000.
XX	PF	02-MAR-2000; 2000WO-US05537.
XX	PR	12-MAR-1999; 99US-0124118.
XX	PA	(REGE-) REGENERON PHARM INC.
XX	P1	Economides AN;
XX	DR	WPI: 2000-638179/61.
XX	DR	P-PSDB; AAB26106.
XX	PT	Novel isolated, human DNA/Cerberus related protein 6 which include
XX	PT	natural homologue, and polypeptides comprising DCR6 domain and nucleic
XX	PS	acids encoding the proteins which are useful as probes and primers -
XX	PS	Claim 2; Fig 3; 40pp; English.
CC	CC	The present sequence comprises exons 1 and 4 of the human
CC	CC	DNA/Cerberus-related protein 6 (hDCR6) coding sequence. It was isolated
CC	CC	from a human kidney cDNA library containing exons 1 and 4 of the
CC	CC	sequence. hDCR6 is closely related to the DAN and DCR5 proteins, both of
CC	CC	which act as antagonists of morphogenic proteins such as BMP. It is
CC	CC	possible that the hDCR6 gene and protein can be used as immunogens,
CC	CC	modulators of cell function, growth and differentiation, to reduce
CC	CC	undesirable bone formation, to identify DCR6 binding agents, in
XX	XX	diagnosis, and in gene therapy.
SO	SO	Sequence 642 BP; 115 A; 230 C; 211 G; 86 T; 0 other;
SO	SO	Query Match 78.9%; Score 501.6; DB 21; Length 642;
SO	SO	Best Local Similarity 87.5%; Pred. No. 8,7e-96;
SO	SO	Matches 562; Conservative 0; Mismatches 74; Indels 6; Gaps
QY	1	ATGACGCCCTCACTAGCCCCGGTGCCTCATCTGTCTACTTTGTGACGCTTCCTTGCTCT
Db	1	ATGCGAGCTCCCACTGGCCTGTGTCTCGCTGCTGCTGTACACAACACCTTCGTGA
QY	61	GTCGAGGCGCAGGGGTGGCGAAGCCTTCAGAAATGATGCCACAGAAGTCAATCCCAGGGCTT
Db	61	GTCGAGGCGCAGGGGTGGCGAAGCCTTCAGAAATGATGCCACAGAAGTCAATCCCAGGGCTT
QY	121	GGAGAGTACCCTCGAGCCTCTCTCC-----TGAAGAACACAGACCATGTAACCGGGCGAG
Db	121	GGAGAGTACCCTCGAGCCTCTCTCTCC-----TGAAGAACACAGACCATGTAACCGGGCGAG
QY	175	AATGAGGCGAAGCTCCCAACCATCTTAATGACGCCMAAGATGTGTCCAGTAGACAGCTGC
Db	175	AACGAGGCGAGGCTCCCAACCATCTTAATGACGCCMAAGATGTGTCCAGTAGACAGCTGC
QY	235	CGCCAGCTGCATAACCCGCTTCTGTACAGACGCGCCATGCGCGCACGATCCCAAGCCGCTC
Db	235	CGCCAGCTGCATAACCCGCTTCTGTACAGACGCGCCATGCGCGCACGATCCCAAGCCGCTC
QY	241	CGCGAGCTGCATTACCCGCTACGTGACCGATGGGCTGTGCCGACGACGACCAAGCCGCTC
Db	241	CGCGAGCTGCATTACCCGCTACGTGACCGATGGGCTGTGCCGACGACGACCAAGCCGCTC
QY	295	ACCGAGTTGGATGTGCTCCGAGCCAGTGGCGGCGCGCGCGCTGTCGCCAAGCCATTCGGG
Db	301	ACCGAGCTGGATGTGCTCCGAGCCAGTGGCGGCGCGCGCGCTGTCGCCAAGCCATTCGGG
QY	355	CGCGTAGATGTGTGCGCGCCGAACGAGACCGGATTTCCGTGTGATCCCGATGCTACCGC
Db	361	CGCGTAGATGTGTGCGCGCCGAACGAGACCGGATTTCCGTGTGATCCCGATGCTACCGC
QY	415	GCGCAGCGGATGTCAGCTGTGTGCCCCGGGGGCGCGCGCGCTCGTGCAGAGGTGCGT
Db	421	GCGCAGCGGATGTCAGCTGTGTGCCCCGGGGTGTGAGGCGCGCGCGCGTGCAGAGGTGCGC
QY	475	CTGGTGGGCTGTGTGAAGTGAAGGAGGCTCAACCGCTTCCAACAACAGTCGAGCTCAAG
Db	481	CTGGTGGGCTGTGTGAAGTGAAGGAGGCTCAACCGCTTCCAACAACAGTCGAGCTCAAG

QY	535	GATTTCGGGCGGAGACCGGCGCGCGGAGAAAGGTCGCAACCGCGGCGCGCGCGG	594		
Db	541	GACTTCGGGAGCCAGGCGCCGCTCGGCGCGAGAGGCGGAGAGCGCGCGCGCGG	600		
QY	595	GGAGCCAAAGCCACACGAGCGGAGCTGAGAAAGCCTACTAG	636		
Db	601	AGGCGCAAAGCCACACGAGCGGCGGAGCTGAGAAAGCCTACTAG	642		
RESULT 6					
AB94293	standard; cDNA; 759 BP.				
XX	AB94293:				
AC	26-MAR-2002	(first entry)			
DT	Human cloaked-2 polypeptide encoding cDNA.				
XX					
DE	Cloaked-2; cysteine knot motif; nephrotropic; cardiact; immunomodulator;				
KM	hepatotropic; antiinflammatory; antihypertoid; cytosatic; neuroprotective;				
KM	antianemic; hypotensive; antiarrhythmic; antidiabetic; muscular;				
KM	antidiabetic; anorectic; gene therapy; cell therapy; antisense therapy;				
KW	human, ss.				
XX					
OS	Homo sapiens.				
XX					
FH	Key	Location/Qualifiers			
FT	CDS	37..678			
FT	/*tag= a	/product= "cloaked-2 polypeptide"			
FT	sig_peptide	37..105			
FT	mat_peptide	106..759			
FT	/*tag= c	/note= "see ABB07207"			
PN	W0200192308-A2.				
XX					
PD	06-DEC-2001.				
XX					
PF	29-MAY-2001; 2001WO-US17478.				
XX					
PR	01-JUN-2000; 2000US-208550P.				
PR	04-AUG-2000; 2000US-223542P.				
XX					
PA	(AMGE-) AMGEN INC.				
XX					
PI	Paszty CJ, Gao Y;				
XX					
DR	WPI; 2002-114325/15.				
XX					
DR	P-PSDB; ABB07207, ABB07209.				
XX					
PT	New human and mouse cysteine-knot polypeptide designated as Cloaked-2,				
XX	for treating or preventing kidney, heart (e.g. myocardial infarction)				
PT	or liver (e.g. hepatitis) diseases -				
XX					
PS	Claim 1; Fig 1; 170p; English.				
XX					
CC	The invention relates to polypeptides comprising a cysteine knot motif				
CC	and designated as Cloaked-2, derived from human and mouse. The cloaked-2				
CC	polypeptides can be expressed by standard recombinant methodology. The				
CC	cloaked-2 polynucleotides are useful in gene therapy and antisense				
CC	therapy. The cloaked-2 polypeptides and polynucleotides are useful for				
CC	treating, preventing, ameliorating or detecting diseases and disorders of				
CC	the kidney (e.g. anemia, hypertension or low blood pressure), heart (e.g.				
CC	cardiac hypertrophy, congestive heart failure, myocardial infarction,				
CC	arrhythmias, atherosclerosis, hypertension or low blood pressure).				
CC	skeletal muscle (e.g. muscular dystrophy or cachexia), placenta (e.g.				
CC	congenital abnormalities or miscarriage), liver (e.g. hepatitis or				
CC	cirrhosis), pancreas (e.g. diabetes or pancreatitis), thyroid (e.g.				
CC	Grave's disease or myxedema) or adrenal cortex (e.g. Cushing's disease				
CC					

or Addison's disease), homeostasis or metabolic diseases (e.g. obesity, cancer or myopathies), infections, or autoimmune diseases. Selective binding agents may be used to modulate the biological activities of cloaked-2 polypeptides or to detect cloaked-2 polypeptide levels in a sample. Transgenic non-human animals are useful for drug candidate screening. The present sequence represents a cDNA encoding the human cloaked-2 polypeptide.

Sequence 759 BP; 125 A; 282 C; 244 G; 108 T; 0 other;

Query Match 78.9%; Score 501.6; DB 24; Length 759;
Best Local Similarity 87.5%; Pred. No. 8.7e-96;
Matches 562; Conservative 0; Mismatches 74; Indels 6; Gaps 1;

```

QY 1 ATGAGCCCTCACTAGCCCGCTGCTCATCTGCTAATTGTGACCGCTTCTGTGCT 60
DB 37 ATGAGCTCCACTGCGCTGTGTCTGTCTGCTGCTGTACACAGCTTCCGTGA 96
QY 61 GTGAGAGGCGCAGGGGTGGCAAGCCTTCAGAAATGATGCCAGAGGTCACTCCAGGGCTT 120
DB 97 GTGAGAGGCGCAGGGGTGGCAAGCCTTCAAGAAATGATGCCAGAGGTCACTCCAGAGCTC 156
QY 121 GGAGAGTACCCCGAGCCTCTCC-----TGAGAACAAACCAAGCAATGAACCGGGCGAG 174
DB 157 GGAGAGTACCCCGAGCCTCCACCGAGCTGAGAACAAACCAAGCAATGAACCGGGCGAG 216
QY 175 AATGAGAGGAGAGCCTCCCAACATCCCTATGAGCGCCAAATGTCCTCCAGTACAGCTGC 234
DB 217 AACGAGAGGCGGCTCCCAACATCCCTTGTAGACCAAAACGTCGCGAGTACAGCTGC 276
QY 235 CGGAGCTGACATACACCCGCTTCTGTACAGAGCGCCATGCGCAGCGCAACCGCGTC 294
DB 277 CGGAGCTGACATACACCCGCTTCTGTACAGAGCGCCATGCGCAGCGCAACCGCGTC 336
QY 295 ACCGAGTTGTGTCTCCGCGAGTGGCGCGCGCGCGCTGTGCTCCCAACCGCATCGCG 354
DB 337 ACCGAGTTGTGTCTCCGCGAGTGGCGCGCGCGCGCTGTGCTCCCAACCGCATCGCG 396
QY 355 CGGCTGAAGTGTGTGGCGCGCGCGCGCGCGCGCGCGCTGTGCTCCCAACCGCATCGCG 414
DB 397 CGGCTGAAGTGTGTGGCGCGCGCGCGCGCGCGCGCGCGCTGTGCTCCCAACCGCATCGCG 456
QY 415 GCGCAGAGGCTGACAGCTGTGTGCGCGCGCGCGCGCGCGCGCGCTGTGCTCCCAACCGCATCGCG 474
DB 457 GCGCAGAGGCTGACAGCTGTGTGCGCGCGCGCGCGCGCGCGCGCGCTGTGCTCCCAACCGCATCGCG 516
QY 475 CTGTGTGCTTGTGTGCAAGTGAAGCGCTTCAACCGCTTCCCAACCAAGTGTGAGCTCAAG 534
DB 517 CTGTGTGCTTGTGTGCAAGTGAAGCGCTTCAACCGCTTCCCAACCAAGTGTGAGCTCAAG 576
QY 535 GACTTGGGCGCGAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 594
DB 577 GACTTGGGCGCGAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 636
QY 595 GAGCCAAAGCCAAACGAGCGGAGCTGAGAACCGCTTCAAG 636
DB 637 AGCGCCAAAGCCAAACGAGCGGAGCTGAGAACCGCTTCAAG 678

```

RESULT 7
AAD27576
ID AAD27576 standard; cDNA; 2271 BP.

AC AAD27576;

DT 18-APR-2002 (first entry)

DE Human osteolevin cDNA.

XX Human, osteolevin; osteopathic; cytostatic; bone formation; osteoporosis;
KW Van Buchem-sclerosteosis disease; sclerosteosis; transgenic animal;
KM Paget's disease; chromosome 17; ss.
XX

OS Homo sapiens.

XX Key Location/Qualifiers

FT CDS 13..654

FT sig_peptide /tag= a

FT mat_peptide /product= "Human osteolevin protein"

FT /tag= b

FT /tag= c

FT /product= "Mature human osteolevin protein"

FT /tag= c

FT /product= "Mature human osteolevin protein"

FT /tag= c

FT /product= "Mature human osteolevin protein"

FT /tag= c

FT /product= "Mature human osteolevin protein"

FT /tag= c

FT /product= "Mature human osteolevin protein"

FT /tag= c

FT /product= "Mature human osteolevin protein"

FT /tag= c

FT /product= "Mature human osteolevin protein"

FT /tag= c

FT /product= "Mature human osteolevin protein"

FT /tag= c

FT /product= "Mature human osteolevin protein"

FT /tag= c

FT /product= "Mature human osteolevin protein"

FT /tag= c

FT /product= "Mature human osteolevin protein"

FT /tag= c

FT /product= "Mature human osteolevin protein"

FT /tag= c

FT /product= "Mature human osteolevin protein"

FT /tag= c

FT /product= "Mature human osteolevin protein"

FT /tag= c

FT /product= "Mature human osteolevin protein"

FT /tag= c

FT /product= "Mature human osteolevin protein"

```

Db 193 AACGAGAGGCGGCTCCCTCCACACCCCTTTGAGACAAAGACGTGCTCGAGTACAGCTGC 252
Qy 235 CGCGAGCTGCACTACACCCGCTTCCTGACAGAGCCCATATGCGCAGCGCAAGCCGCTC 294
Db 253 CGCGAGCTGCACTACACCCGCTTCCTGACAGAGCCCATATGCGCAGCGCAAGCCGCTC 312
Qy 295 ACCGAGTTGTTGCTCTCGGCGCAGTGGCGCCCGCGCGCTGCTGCCAAGCCCATCGG 354
Db 313 ACCGAGTTGTTGCTCTCGGCGCAGTGGCGCCCGCGCGCTGCTGCCAAGCCCATCGG 372
Qy 355 CGCGTGAAGTGTGGCGCCCGCAAGACCGGATTTCCGCTGCACTCCGATCCGCTACCG 414
Db 373 CGCGCAAGTGTGGCGCCCGCAAGTGGCGCCCGCAACTTCCGCTGCACTCCGATCCGCTACCG 432
Qy 415 GCGCAGCGGCTGCACTGCTGTGCGCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 474
Db 433 GCGCAGCGGCTGCACTGCTGTGCGCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 492
Qy 475 CTGCTGCGCTTGTGCAAGTGCAGAGCGCTTCAACCGCTTCAACAGTGCAGTCAAG 534
Db 493 CTGCTGCGCTTGTGCAAGTGCAGAGCGCTTCAACCGCTTCAACAGTGCAGTCAAG 552
Qy 535 GACTTCCGCGCGCGGAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 594
Db 553 GACTTCCGCGCGCGGAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 612
Qy 595 GAGCGCAAGCGCAACGAGCGGAGCTGAGAGCGCTTACG 636
Db 613 AGCGCCAAAGCGCAACGAGCGGAGCTGAGAGCGCTTACG 654

RESULT 8
AAA29055
ID AAA29055 standard; cDNA; 2301 BP.
XX
AC AAA29055;
XX
DT 12-SEP-2000 (first entry)
XX
DE Human TGF-beta binding protein (BBER) cDNA.
XX
KM osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
KW BBER; chromosome 17q12-21; gene therapy; antisense therapy; fracture;
XX bone mineralization; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 48..689
FT /tag=a
FT /product=TGF-beta_binding_protein
XX
PN MO200032773-A1.
XX
PD 08-JUN-2000.
XX
PF 24-NOV-1999; 99WO-US27990.
XX
PR 27-NOV-1998; 98US-0110283.
XX
PA (DARW-) DARWIN DISCOVERY LTD.
XX
PI Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepers BW,
XX Van Nese J, Winkler DG;
XX
DR MPI: 2000-41231/35.
XX P-PSDB; AA196429.
XX
PT Nucleic acids (1) encoding a transforming growth factor beta binding
PT protein, useful for identifying agents for treating osteopenia,
PT osteoporosis and fractures
XX

```

```

PS Claim 1; Page 114-115; 162pp; English.
XX
CC This cDNA encodes a human transforming growth factor-beta (TGF-beta)
CC binding protein designated BBER. The hBER gene has been localized
CC to the chromosome 17q12-21. The cDNA and protein may be used for
CC prevention, treatment and diagnosis of diseases associated with
CC inappropriate BBER expression. For example, they may be used to treat
CC disorders associated with decreased TGF-beta BP expression. The cDNA or
CC vectors may be administered to treat diseases by rectifying mutations or
CC deletions in a patient's genome that affect the activity of BBER by
CC expressing inactive proteins or to supplement the patients own production
CC of BBER polypeptides. The nucleic acids may be used for recombinant
CC production of BBER, gene therapy, antisense therapy, as probes for
CC diagnostic assays and for functional studies. BBER may be used to raise
CC antibodies and for identification of BBER modulators. BBER antagonists
CC may be used to increase bone mineral content for the treatment of
CC disorders such as osteopenia, osteoporosis, fractures and other
CC disorders associated with low mineral content.
XX
SQ Sequence 2301 BP; 568 A; 635 C; 614 G; 484 T; 0 other;
XX
Query Match 78.9%; Score 501.6; DB 21; Length 2301;
Best Local Similarity 87.5%; Pred. No. 9.2e-96;
Matches 562; Conservative 0; Mismatches 74; Indels 6; Gaps 1;
Qy 1 ATGCAAGCTCTCACTAGCCCGGCTCATCTGCTTCTGACAGCGCTTCTGTCT 60
Db 48 ATGCAAGCTCTCACTAGCCCGGCTCATCTGCTTCTGACAGCGCTTCTGTCT 107
Qy 61 GTGAGAGGCGCAAGGCGTGGCAAGCTTCAAGAAATGATCCACAGAGTCAATCCAGAGCTT 120
Db 108 GTGAGAGGCGCAAGGCGTGGCAAGCTTCAAGAAATGATCCACAGAGTCAATCCAGAGCTT 167
Qy 121 GGAAGATACCCCGAGCTCTCTCC-----TGAGAACACCAAGCATGAAACCGGCGGAG 174
Db 168 GGAAGATACCCCGAGCTCTCTCC-----TGAGAACACCAAGCATGAAACCGGCGGAG 227
Qy 175 AATGAGAGCAAGCTCTCCCAACCATCTATGACGCGCAAGATGTCTCGAGTACAGCTGC 234
Db 228 AATGAGAGCGGCGCTCTCCCAACCATCTATGAGCAAGAGTGTCCAGTACAGCTGC 287
Qy 235 CGCGAGCTGCACTACACCGCTTCTGACAGCGCGCATGCGCAGCGCGCGCGCTC 294
Db 288 CGCGAGCTGCACTACACCGCTTCTGACAGCGCGCATGCGCGCGCGCGCGCGCGCTC 347
Qy 295 ACCGAGTTGTTGCTCTCGGCGCAGTGGCGCCCGCGCGCGCTGCGCGCAATCGGCTCGGG 354
Db 348 ACCGAGTTGTTGCTCTCGGCGCAGTGGCGCCCGCGCGCGCTGCGCGCAATCGGCTCGGG 407
Qy 355 CGCGTGAAGTGTGGCGCGCGCGCAAGCGGATTTCCGCTGATCCGGAATGCTACCGC 414
Db 408 CGCGTGAAGTGTGGCGCGCGCGCAAGCGGATTTCCGCTGATCCGGAATGCTACCGC 467
Qy 415 GCGCAGCGGCTGCACTGCTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 474
Db 468 GCGCAGCGGCTGCACTGCTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 527
Qy 475 CTGCTGCGCTTGTGCAAGTGCAGAGCGCTCAACCGCTTCAACAGTGCAGTCAAG 534
Db 528 CTGCTGCGCTTGTGCAAGTGCAGAGCGCTCAACCGCTTCAACAGTGCAGTCAAG 587
Qy 535 GACTTCCGCGCGGAGACCGCGCGCGCGCAAGAGGTTCGAAGCGCGCGCGCGCGCGCGCG 594
Db 588 GACTTCCGCGCGGAGACCGCGCGCGCGCAAGAGGTTCGAAGCGCGCGCGCGCGCGCGCG 647
Qy 595 GAGCGCAAGCGCAACGAGCGGAGCTGAGAAACGCTTACTG 636
Db 648 AGCGCCAAAGCGCAACGAGCGGAGCTGAGAAACGCTTACTG 689

RESULT 9
AAA29056
ID AAA29056 standard; cDNA; 2301 BP.

```

XX AC AAA29056;
 XX 12-SEP-2000 (first entry)
 XX DE Human TGF-beta binding protein (BEER) variant V101 cDNA.
 XX osteoplastic; transforming growth factor-beta; TGF-beta; binding protein;
 XX BEER; variant; V101; gene therapy; antisense therapy; fracture;
 XX chromosome 17q12-21; bone mineralization; ss.
 XX OS Homo sapiens.
 XX FH Key Location/Qualifiers
 XX FT CDS 48..689
 XX FT /tag= a
 XX FT /label= BEER variant V101
 XX FT /product= TGF-beta_binding_protein
 XX PN MO200032773-A1.
 XX 08-JUN-2000.
 XX PF 24-NOV-1999; 99WO-US27990.
 XX PR 27-NOV-1998; 98US-0110283.
 XX PA (DARW-) DARWIN DISCOVERY LTD.
 XX PI Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepert BW,
 XX PI Van Ness J, Winkler DG;
 XX DR WPI: 2000-412321/35.
 XX DR P-PSDB; AAY96430.
 XX PT Nucleic acids (I) encoding a transforming growth factor beta binding
 XX PT protein, useful for identifying agents for treating osteopenia,
 XX PT osteoporosis and fractures
 XX PS Claim 1; Page 118-119; 162pp; English.
 XX CC This cDNA encodes a variant human transforming growth factor-beta
 XX CC (TGF-beta) binding protein designated BEER V101. The encoded protein
 XX CC comprises a substitution of isoleucine for the wild-type valine at
 XX CC residue 10. The cDNA and protein may be used for prevention, treatment,
 XX CC and diagnosis of diseases associated with inappropriate BEER expression.
 XX CC For example, they may be used to treat disorders associated with
 XX CC decreased TGF-beta BP expression. The cDNA or vectors may be administered
 XX CC to treat diseases by rectifying mutations or deletions in a patient's
 XX CC genome that affect the activity of BEER by expressing inactive proteins
 XX CC or to supplement the patient's own production of BEER polypeptides. The
 XX CC nucleic acids may be used for recombinant production of BEER. Gene
 XX CC therapy, antisense therapy, as probes for diagnostic assays and for
 XX CC functional studies. BEER may be used to raise antibodies and for
 XX CC identification of BEER modulators. BEER antagonists may be used to
 XX CC increase bone mineral content for the treatment of disorders such as
 XX CC osteopenia, osteoporosis, fractures and other disorders associated with
 XX CC low mineral content.
 XX SQ Sequence 2301 BP; 569 A; 634 C; 614 G; 484 T; 0 other;
 XX
 XX Query Match 78.9%; Score 501.6; DB 21; Length 2301;
 XX Best Local Similarity 87.5%; Pred. No. 9.2e-96;
 XX Matches 562; Conservative 0; Mismatches 74; Indels 6; Gaps 1;
 XX
 XX QY 1 ATGAGAGCCCTCATTAGCCCGCTGCTCATCTGCTACTTGTGACGCTGCTTCTGTGCT 60
 XX DB 48 ATGAGAGCTCCACATGCGCTGTGTCTCATCTGCTGTGTAACAACAAGCTTCCGCTGA 107
 XX QY 61 GTGAGAGGCCAGGGGTGACCACTTCAAGAAATGATGCCACAGAGTCAATCCAGGAGCTT 120
 XX DB 108 GTGAGAGGCCAGGGGTGACGCTTCAAGAAATGATGCCACAGAAATCAATCCGAGAGCTC 167

QY 121 GGAGAGTACCCCGAGCCTCTCC-----TGAGAACACCAACCAATGAACCGGCGGAG 174
 DB 168 GGAGAGTACCCCGAGCCTCTCCACCGAGCTGAGAGAACAAAGACATGAACCGGCGGAG 227
 QY 175 AATGAGGCGAGACTCTCCACCATCTTATGACGCCAAAGTGTGTCCGATCAAGTTC 234
 DB 228 AACGAGGCGGCGCTCTCCACCATCTTATGACGCCAAAGTGTGTCCGATCAAGTTC 287
 QY 235 CGCGAGCTGACCTACACCGGCTCTGACAGACGAGCCATATGCGGACGCGCAAGCGGCT 294
 DB 288 CGCGAGCTGACCTACACCGGCTCTGACAGACGAGCCATATGCGGACGCGCAAGCGGCT 347
 QY 295 ACCGAGTTGATGCTCTCCGAGCAAGTCCGCGGCGGCTCTGCTCCAAAGCCATCGG 354
 DB 348 ACCGAGCTGATGCTCTCCGAGCAAGTCCGCGGCGGCTCTGCTCCAAAGCCATCGG 407
 QY 355 CGCGTGAAGTGTGCGCGCGGCAAGACGAGATTTCCGCTGCAATCCCGATCCGTACCG 414
 DB 408 CGCGGCAAGTGTGTGCGCGGCAAGTGTGCGGCGGATTTCCGCTGCAATCCCGATCCCG 467
 QY 415 GCGCAGCGGCGTGTGAGCTGTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 474
 DB 468 GCGCAGCGGCGTGTGAGCTGTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 527
 QY 475 CTGTGAGCCTGTGTCAAGTGTCAAGCGGCTCAACCGCTTCAACAACAGTGTGAGTCAAG 534
 DB 528 CTGTGAGCCTGTGTCAAGTGTGTCAAGCGGCTCAACCGCTTCAACAACAGTGTGAGTCAAG 587
 QY 535 GACTTGTGCGCGGAGACCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 594
 DB 588 GACTTGTGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 647
 QY 595 GGAGCCAAAGCCAAACAGGCGGAGCTGAGAACGCTTACTAG 636
 DB 648 AGCGCCAAAGCCAAACAGGCGGAGCTGAGAACGCTTACTAG 689
 XX
 XX RESULT 10
 XX ID AAA91023
 XX ID AAA91023 standard; DNA; 2329 BP.
 XX AC AAA91023;
 XX DT 05-APR-2001 (first entry)
 XX DE Human secreted protein PRO7476 coding sequence.
 XX KW Secreted protein; human; PRO protein; neoplastic cell growth; tumour;
 XX KW proliferation; leukaemia; lymphoid malignancy; inflammatory disorder;
 XX KW angiogenic disorder; immunologic disorder; PRO7476; ds.
 XX OS Homo sapiens.
 XX FH Key Location/Qualifiers
 XX FT CDS 62..703
 XX FT /tag= a
 XX FT /product= PRO7476
 XX PN MO200075317-A2.
 XX PD 14-DEC-2000.
 XX PF 15-MAY-2000; 2000WO-US13358.
 XX PR 09-JUN-1999; 99US-0138385.
 XX PR 20-JUN-1999; 99US-0144790.
 XX PR 03-AUG-1999; 99US-0146843.
 XX PR 10-AUG-1999; 99US-0148188.
 XX PR 17-AUG-1999; 99US-0149320.
 XX PR 17-AUG-1999; 99US-0149327.
 XX PR 17-AUG-1999; 99US-0149396.
 XX PR 20-AUG-1999; 99US-0150114.
 XX PR 31-AUG-1999; 99US-0151700.

Db 108 GTGAGAGGCTAGAGGAGTGGCAGAGCGCTTCAGAGATGATGCCAGAAATCATCCCCGAGCTC 167
 Qy 121 GGAAGATACCCCGAGAGCTCTCTCC-----TGAGAACCAACGAGCACTGAAACGGGGGAG 174
 Db 168 GGAAGATACCCCGAGAGCTCTCTCC-----TGAGAACCAACGAGCACTGAAACGGGGGAG 227
 Qy 175 AATGAGAGGAGAGCTCTCTCCACCACTCCCTATGAGCGCCAGAGATGTGTCCAGTACAGCTGC 234
 Db 228 AACGAGAGGAGGAGCTCTCTCCACCACTCCCTATGAGCGCCAGAGATGTGTCCAGTACAGCTGC 287
 Qy 235 CGGAGAGTGCATCAACACCGCTCTCTGACAGAGAGGCCCATATGCCGACGCGCATC 294
 Db 288 CGGAGAGTGCATCAACACCGCTCTCTGACAGAGAGGCCCATATGCCGACGCGCATC 347
 Qy 295 ACCGAGTGTGTGTCTCTCGGCGCAGTGGCGCCCGCGCGCTGTGTCGCCAACGCGCATCGAG 354
 Db 348 ACCGAGTGTGTGTCTCTCGGCGCAGTGGCGCCCGCGCGCTGTGTCGCCAACGCGCATCGAG 407
 Qy 355 CGCGTGAAGTGTGTGTCTCTCGGCGCAGTGGCGCCCGCGCGCTGTGTCGCCAACGCGCATCGAG 414
 Db 408 CGCGGCAAGTGTGTGTCTCTCGGCGCAGTGGCGCCCGCGCGCTGTGTCGCCAACGCGCATCGAG 467
 Qy 415 GCGCAGAGGAGTGCAGTGTGTGTCTCTCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 474
 Db 468 GCGCAGAGGAGTGCAGTGTGTGTCTCTCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 527
 Qy 475 CTGAGTGGCTCTGTGCAAGTGCAGAGCGCTCTGACACCGCTCTGACACCGAGTGGAGCTCAAG 534
 Db 528 CTGAGTGGCTCTGTGCAAGTGCAGAGCGCTCTGACACCGCTCTGACACCGAGTGGAGCTCAAG 587
 Qy 535 GACTTGGGCGCGGAGACCG 594
 Db 588 GACTTGGGCGCGGAGACCG 647
 Qy 595 GAGGCCAAGCCAAACCAACGAGCGAGCTGAGAGAGCTTACTAG 636
 Db 648 AGCGCCAAAGCCAAACCAACGAGCGAGCTGAGAGAGCTTACTAG 689
 RESULT 13
 AAA29062
 ID AAA29062 standard; cDNA, 2301 BP.
 XX
 AC AAA29062;
 XX
 DT 12-SEP-2000 (first entry)
 XX
 DE Human TGF-beta binding protein (BBER) variant P38R cDNA.
 XX
 KW osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
 KM BBER; variant; P38R; gene therapy; antisense therapy; fracture;
 KM chromosome 17q12-21; bone mineralization; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 48..689
 FT /tag= a
 FT /label= BBER variant_P38R
 FT /product= TGF-beta_binding_protein
 XX
 PN WO200032773-A1.
 PD 08-JUN-2000.
 PF 24-NOV-1999; 99WO-US27990.
 XX
 PR 27-NOV-1998; 98US-0110283.
 XX
 PA (DARW-) DARWIN DISCOVERY LTD.
 XX
 PI Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepker BW,
 Van Ness J, Winkler DG;

XX WPI: 2000-412321/35.
 DR P-PSDB; AAY96436#.
 XX
 PT Nucleic acid (1) encoding a transforming growth factor beta binding
 PT protein, useful for identifying agents for treating osteopenia,
 PT osteoporosis and fractures
 XX
 PS Claim 27; Page 120-121; 162pp; English.
 XX
 CC This cDNA encodes a variant human transforming growth factor-beta
 CC (TGF-beta) binding protein designated BBER P38R. The encoded protein
 CC comprises a substitution of arginine for the wild-type proline at
 CC residue 38. The cDNA and protein may be used for prevention, treatment,
 CC and diagnosis of diseases associated with inappropriate BBER expression.
 CC For example, they may be used to treat disorders associated with
 CC decreased TGF-beta BP expression. The cDNA or vectors may be administered
 CC to treat diseases by rectifying mutations or deletions in a patient's
 CC genome that affect the activity of BBER by expressing inactive proteins
 CC or to supplement the patient's own production of BBER polypeptides. The
 CC nucleic acids may be used for recombinant production of BBER, gene
 CC therapy, antisense therapy, as probes for diagnostic assays and for
 CC functional studies. BBER may be used to raise antibodies and for
 CC identification of BBER modulators. BBER antagonists may be used to
 CC increase bone mineral content for the treatment of disorders such as
 CC osteopenia, osteoporosis, fractures and other disorders associated with
 CC low mineral content.
 XX
 SQ Sequence 2301 BP; 568 A; 634 C; 615 G; 484 T; 0 other;
 Query Match 78.6%; Score 500; DB 21; Length 2301;
 Best Local Similarity 87.4%; Pred. No. 2e-95;
 Matches 561; Conservative 0; Mismatches 75; Indels 6; Gaps 1;
 Qy 1 ATGCAAGCCCTCACTAGCCCGCTGCTCTATCTGCTTGTGACGCTGCTCTGTGCT 60
 Db 48 ATGCAAGCTCCCACTGAGCCCTGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 107
 Qy 61 GTGAGAGGCGCAGGGGTGTGCAAGCTTTCAGGAATATGTCACAGAGGTATCCAGGGCTT 120
 Db 108 GTGAGAGGCGCAGGGGTGTGCAAGCTTTCAGGAATATGTCACAGGAATATATCCGAGCTTC 167
 Qy 121 GGAAGATACCCCGAGAGCTCTCTCC-----TGAGAACCAACGAGCACTGAAACGGGGGAG 174
 Db 168 GGAAGATACCCCGAGAGCTCTCTCC-----TGAGAACCAACGAGCACTGAAACGGGGGAG 227
 Qy 175 AATGAGAGGAGAGCTCTCTCCACCACTCCCTATGAGCGCCAGAGATGTGTCCAGTACAGCTGC 234
 Db 228 AACGAGAGGAGGAGCTCTCTCCACCACTCCCTATGAGCGCCAGAGATGTGTCCAGTACAGCTGC 287
 Qy 235 CGGAGAGTGCATCAACACCGCTCTCTGACAGAGAGGCCCATATGCCGACGCGCATC 294
 Db 288 CGGAGAGTGCATCAACACCGCTCTCTGACAGAGAGGCCCATATGCCGACGCGCATC 347
 Qy 295 ACCGAGTGTGTGTCTCTCGGCGCAGTGGCGCCCGCGCGCTGTGTCGCCAACGCGCATCGAG 354
 Db 348 ACCGAGTGTGTGTCTCTCGGCGCAGTGGCGCCCGCGCGCTGTGTCGCCAACGCGCATCGAG 407
 Qy 355 CGCGTGAAGTGTGTGTCTCTCGGCGCAGTGGCGCCCGCGCGCTGTGTCGCCAACGCGCATCGAG 414
 Db 408 CGCGGCAAGTGTGTGTCTCTCGGCGCAGTGGCGCCCGCGCGCTGTGTCGCCAACGCGCATCGAG 467
 Qy 415 GCGCAGAGGAGTGCAGTGTGTGTCTCTCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 474
 Db 468 GCGCAGAGGAGTGCAGTGTGTGTCTCTCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 527
 Qy 475 CTGAGTGGCTCTGTGCAAGTGCAGAGCGCTCTGACACCGCTCTGACACCGAGTGGAGCTCAAG 534
 Db 528 CTGAGTGGCTCTGTGCAAGTGCAGAGCGCTCTGACACCGCTCTGACACCGAGTGGAGCTCAAG 587
 Qy 535 GACTTGGGCGCGGAGACCG 594
 Db 588 GACTTGGGCGCGGAGACCG 647

QY 595 GGAGCCAAAGCAACGAGCGGAGCTTGAGAGCCCTACTAG 636
 DB 648 AGCGCCAAAGCAACGAGCGGAGCTTGAGAGCCCTACTAG 689

RESULT 14

AAA29063
 ID AAA29063 standard; DNA; 35828 BP.

AC AAA29063;

DT 12-SEP-2000 (first entry)

DE Murine TGF-beta binding protein (BBER) genomic MluI-AvII DNA fragment.

KW Osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
 BBER; gene therapy; antisense therapy; fracture; bone mineralization;
 XX transgenic mouse; over-expression; ds.

OS Mus musculus.

XX Key Location/Qualifiers

FT CDS 18317..21464

FT /*tag= a

FT 18317..18530

FT /*tag= b

FT /*tag= c

FT 18531..21022

FT /*tag= d

FT 21023..21464

FT /*tag= d

FT /*tag= d

FT /*tag= d

FT /*tag= d

FT /*tag= d

FT /*tag= d

FT /*tag= d

FT /*tag= d

FT /*tag= d

FT /*tag= d

FT /*tag= d

FT /*tag= d

FT /*tag= d

FT /*tag= d

FT /*tag= d

FT /*tag= d

FT /*tag= d

FT /*tag= d

FT /*tag= d

FT /*tag= d

FT /*tag= d

FT /*tag= d

FT /*tag= d

FT /*tag= d

FT /*tag= d

FT /*tag= d

FT /*tag= d

FT /*tag= d

FT /*tag= d

FT /*tag= d

FT /*tag= d

FT /*tag= d

FT /*tag= d

FT /*tag= d

FT /*tag= d

FT /*tag= d

FT /*tag= d

FT /*tag= d

FT /*tag= d

FT /*tag= d

FT /*tag= d

FT /*tag= d

FT /*tag= d

FT /*tag= d

XX SQ Sequence 35828 BP; 8973 A; 8817 C; 9064 G; 8973 T; 1 other;

Query Match 67.1%; Score 426.6; DB 21; Length 35828;

Best Local Similarity 95.8%; Pred. No. 4; se-80;

Matches 438; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 180 AGGAGACCTCCCAACATCCCTATGACGCCCAAGATGTGTCCGATGACGCTCCCGCA 239

DB 20988 AGGCTGACAGCCCTTCACGATCCTCTCTCCGACATGTGTCCGATGACGCTCCCGCA 21047

QY 240 GCTGACATACACCCGCTTCTGACAGAGCCCATGCGCGCAAGCCGCTACCGA 299

DB 21048 GCTGACATACACCCGCTTCTGACAGAGCCCATGCGCGCAAGCCGCTACCGA 21107

QY 300 GTTGTGTGTCTCCGCGCGAGCCGCGCGCGCGCTCTGCTCCCAAGCCCATGCGCGCG 359

DB 21108 GTTGTGTGTCTCCGCGCGAGCCGCGCGCGCGCTCTGCTCCCAAGCCCATGCGCGCG 21167

QY 360 GAAAGTGTGTGCGCGCGCGCAACCGGATTTCCGCTGATCCCGGATGCTACCGCGCGCA 419

DB 21168 GAAAGTGTGTGCGCGCGCGCAACCGGATTTCCGCTGATCCCGGATGCTACCGCGCGCA 21227

QY 420 GCGGCTGTGACGCTGTGCTGCTCCCGCGCGCGCGCGCTGCGCAAGGTGCTTGTGT 479

DB 21228 GCGGCTGTGACGCTGTGCTGCTCCCGCGCGCGCGCGCTGCGCAAGGTGCTTGTGT 21287

QY 480 GGGCTGTGTGCAAGTGTGCAAGCGGCTCACCGGCTTCCACACACGATGCGAGTCAAGGACTT 539

DB 21288 GGGCTGTGTGCAAGTGTGCAAGCGGCTCACCGGCTTCCACACACGATGCGAGTCAAGGACTT 21347

QY 540 CGGCGCGGAGACCGCGCGCGCGCGCAAGAGGTGTGCAAGCGGCGCGCGCGCGCGCGCGCG 599

DB 21348 CGGCGCGGAGACCGCGCGCGCGCGCAAGAGGTGTGCAAGCGGCGCGCGCGCGCGCGCGCG 21407

QY 600 CAAAGCCAAACGAGCGGAGCTTGAGAGCCCTACTAG 636

DB 21408 CAAAGCCAAACGAGCGGAGCTTGAGAGCCCTACTAG 21444

RESULT 15

AAA29060
 ID AAA29060 standard; CDNA; 532 BP.

AC AAA29060;

DT 12-SEP-2000 (first entry)

DE Bovine TGF-beta binding protein (BBER) cDNA.

KW osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
 BBER; gene therapy; antisense therapy; fracture; bone mineralization; ss.

OS Bos taurus.

XX Key Location/Qualifiers

FT CDS 1..530

FT /*tag= a

FT /*tag= a

FT /*tag= a

FT /*tag= a

FT /*tag= a

FT /*tag= a

FT /*tag= a

FT /*tag= a

FT /*tag= a

FT /*tag= a

FT /*tag= a

FT /*tag= a

FT /*tag= a

FT /*tag= a

FT /*tag= a

FT /*tag= a

FT /*tag= a

FT /*tag= a

FT /*tag= a

FT /*tag= a

PI Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepker BW,
 Van Ness J, Winkler DG;

Example 9; Page 74-87; 162pp; English.
 This 35 kb MluI-AvII genomic DNA restriction fragment includes the
 entire murine transforming growth factor-beta (TGF-beta) binding protein
 (BBER) gene. This DNA was used to generate a transgenic mouse
 over-expressing the BBER gene. The gene and protein may be used for
 prevention, treatment and diagnosis of diseases associated with
 inappropriate BBER expression. For example, they may be used to treat
 disorders associated with decreased TGF-beta BP expression. The cDNA or
 vectors may be administered to treat diseases by rectifying mutations or
 deletions in a patient's genome that affect the activity of BBER by
 expressing inactive proteins or to supplement the patients own production
 of BBER polypeptides. The nucleic acids may be used for recombinant
 production of BBER, gene therapy, antisense therapy, as probes for
 diagnostic assays and for functional studies. BBER may be used to raise
 antibodies and for identification of BBER modulators. BBER antagonists
 may be used to increase bone mineral content for the treatment of
 disorders such as osteopenia, osteoporosis, fractures and other disorders
 associated with low mineral content.

XX
DR WPI: 2000-412321/35.
DR P-PSDB; AAY96434.

XX Nucleic acids (1) encoding a transforming growth factor beta binding
PT protein, useful for identifying agents for treating osteopenia,
PT osteoporosis and fractures

XX
PS Claim 1, Page 126, 162pp; English.

XX
CC This CDNA encodes a bovine transforming growth factor-beta (TGF-beta)
CC binding protein designated bBFR. The CDNA and protein may be used for
CC prevention, treatment and diagnosis of diseases associated with
CC inappropriate BFR expression. For example, they may be used to treat
CC disorders associated with decreased TGF-beta BP expression. The CDNA or
CC vectors may be administered to treat diseases by rectifying mutations or
CC deletions in a patient's genome that affect the activity of BFR by
CC expressing inactive proteins or to supplement the patient's own production
CC of BFR polypeptides. The nucleic acids may be used for recombinant
CC production of BFR, gene therapy, antisense therapy, as probes for
CC diagnostic assays and for functional studies. BFR may be used to raise
CC antibodies and for identification of BFR modulators. BFR antagonists
CC may be used to increase bone mineral content for the treatment of
CC disorders such as osteopenia, osteoporosis, fractures and other disorders
CC associated with low mineral content.

XX
SQ Sequence 532 BP, 95 A; 199 C; 176 G; 62 T; 0 other;

Query Match 65.4%; Score 415.8; DB 21; Length 532;

Best Local Similarity 87.8%; Pred. No. 6, 7e-78;

Matches 466; Conservative 0; Mismatches 62; Indels 3; Gaps 1;

QY 90 GAATGATGCCACAGAGTATCCAGAGGCTTGGAGATACCCGAGCCTCTCTCTGA--- 146
DB 2 GAATGATGCCACAGAAATCATCCCGAGCTGGGCGAGTACCCGAGCCTCTGCGAGAGCT 61
QY 147 GAACAAACAGACCATGAACCGGGGGGAGAAATGAGGAGACACTCCCAACATCCCTATGA 206
DB 62 GAACAAACAGACCATGAACCGGGGGGAGAAATGAGGAGACACTCCCAACATCCCTATGA 121
QY 207 CGCCAAAGATGTGTCAGATACAGTCCCGAGACTGACATACACCGCTTCTCTGACAGA 266
DB 122 GACCAAGAGAGCCTCCAGATACAGTCCCGAGACTGACATACACCGCTTCTCTGACAGA 181
QY 267 CGGCCATGCCGAGCGCCCAAGCCGATCACCGAGTTGTGCTCCGACCAAGTCCGCC 326
DB 182 TGGGCGGTGCGGAGCGCCCAAGCCGATCACCGAGTTGTGCTCCGACCAAGTCCGCC 241
QY 327 CGCGCGGTGTCGCGCAAGCGCCCAAGCCGATCACCGAGTTGTGCTCCGACCAAGTCCGCC 386
DB 242 GGGCGGCTCTGTCGCAAGCGCCCAAGCGCCCAAGTGTGGCGCCCAAGCGGCGCCGA 301
QY 387 TTTCCGCTGATCCCGGATGCTACCGCGGAGCGGGTGACGCTGTGTGCTCCCGGGGG 446
DB 302 CTTCGCTGATCCCGGATGCTACCGCGGAGCGGGTGACGCTGTGTGCTCCCGGGGG 361
QY 447 CGCGGCGCTGCTGCGGCAAGGTGCTGTGTGCTGTGCAAGTGAAGCGCTCTAC 506
DB 362 CGCGGCGCTGCTGCGGCAAGGTGCTGTGTGCTGTGCAAGTGAAGCGCTCTAC 421
QY 507 CCGTTTCAACAACAGTGAAGTCAAGGACTTGGGCGCGAGACCGGCGCGGAGAA 566
DB 422 TCGCTTCAACAACAGTGAAGTCAAGGACTTGGGCGCGAGACCGGCGCGGAGAA 481
QY 567 GGGTGCAGACCGGCGCGCGGAGGAGCAAGCAACCAAGGCGAG 617
DB 482 GGGCGGAGAGTGTGGCGCGCGGAGGAGCAAGCAACCAAGGCGAG 532

Search completed: March 28, 2003, 23:23:25
Job time : 215.234 secs

•
•
•
•

0

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 29, 2003, 00:44:05 ; Search time 2065.75 Seconds

(without alignments)
8960.144 Million cell updates/sec

Title: US-09-867-274-3

Perfect score: 636
Sequence: 1 atcgagccctcactagcccc.....agctgagagcgcctactag 636Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 2054640 seqs, 14551402878 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

```
1: gb_da:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pac:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_str:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_da:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_str:*
28: em_un:*
29: em_vl:*
30: em_hcg_hum:*
31: em_hcg_inv:*
32: em_hcg_other:*
33: em_hcg_mus:*
34: em_hcg_pln:*
35: em_hcg_rtd:*
36: em_hcg_mam:*
37: em_hcg_vtl:*
38: em_sy:*
39: em_hcg_hum:*
40: em_hcg_mus:*
41: em_hcg_other:*
```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	636	100.0	636	6 AX323455	AX323455 Sequence
2	585	92.0	638	10 AF326740	AF326740 Mus muscu
3	424	66.7	81806	10 AF326737	AF326737 Mus muscu
4	424	66.7	110000	2 AC068782-2	AC068782-2
5	424	66.7	205277	2 AC012296	Continuation (3 of
6	424	66.7	208135	10 AC068807	AC068807 Mus muscu
7	373	58.6	198508	10 AL591145	AL591145 Mouse DNA
8	214	33.6	51575	2 AC023810	AC023810 Mus muscu
9	104	16.4	674	10 AF326741	AF326741 Rattus no
10	104	16.4	101804	2 AC098160	AC098160 Rattus no
11	104	16.4	104898	2 AC121721	AC121721 Rattus no
12	69	10.8	759	6 AX323453	AX323453 Sequence
13	69	10.8	2271	6 AX342535	AX342535 Sequence
14	69	10.8	2296	9 AF331844	AF331844 Homo sapi
15	69	10.8	2323	9 AF326739	AF326739 Homo sapi
16	69	10.8	2329	6 AX056687	AX056687 Sequence
17	69	10.8	7099	6 AX342537	AX342537 Sequence
18	69	10.8	21501	9 AF326736	AF326736 Homo sapi
19	69	10.8	80117	9 AC055813	AC055813 Homo sapi
20	69	10.8	94752	9 AC003098	AC003098 Homo sapi
21	69	10.8	177744	2 AC073954	AC073954 Homo sapi
22	68	10.7	642	9 AF326742	AF326742 Cercopit
23	32	5.0	532	4 AF326738	AF326738 Bos tauru
24	26	4.1	41	6 AX323472	AX323472 Sequence
25	25	3.9	25	6 AX323466	AX323466 Sequence
26	25	3.9	45	6 AX323471	AX323471 Sequence
27	24	3.8	24	6 AX323464	AX323464 Sequence
28	21	3.3	21	6 AX323467	AX323467 Sequence
29	21	3.3	21	6 AX323469	AX323469 Sequence
30	21	3.3	27	6 AX056700	AX056700 Sequence
31	21	3.3	29	6 AX323461	AX323461 Sequence
32	21	3.3	3552	9 AB002058	AB002058 Homo sapi
33	21	3.3	28984	9 AB002059	AB002059 Homo sapi
34	21	3.3	147086	9 AC002472	AC002472 Homo sapi
35	21	3.3	152172	8 AC068924	AC068924 Oryza sat
36	21	3.3	162470	9 AC007664	AC007664 Homo sapi
37	20	3.1	24	6 AX323462	AX323462 Sequence
38	20	3.1	10460	1 AE011791	AE011791 Xanthomon
39	20	3.1	14108	1 AE004989	AE004989 Halobacte
40	20	3.1	45653	2 AC105239	AC105239 Homo sapi
41	20	3.1	69337	2 AC101261	AC101261 Mus muscu
42	20	3.1	97216	2 AC121752	AC121752 Rattus no
43	20	3.1	97684	9 HS309122	HS309122 Human DNA
44	20	3.1	136081	2 AC128778	AC128778 Rattus no
45	20	3.1	139998	2 AC111270	AC111270 Rattus no

ALIGNMENTS

RESULT 1
AX323455 LOCUS AX323455 636 bp DNA linear PAT 07-JAN-2002
DEFINITION Sequence 3 from Patent WO01923308.
AX323455
VERSION AX323455.1 GI:18094217
KEYWORDS
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS Paszty, C.J. and Gao, Y.
TITLE Cystine-knot polypeptides: cloaked-2 molecules and uses thereof
JOURNAL Patent: WO 01923308-A 3 06-DEC-2001;

Amgen, Inc. (US)
Location/Qualifiers
1. .636
/organism="Mus musculus"
/db_xref="taxon:10090"

BASE COUNT 114 a 224 c 207 g 91 t
ORIGIN

Query Match 100.0%; Score 636; DB 6; Length 636;
Best Local Similarity 100.0%; Pred. No. 8.3e-305;
Matches 636; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGAGCCCTCACTAGCCGCTGCTCTATCTGCTTCTTGTGACAGCTGCTTCTTGTCT 60
DB 1 ATGAGAGCCCTCACTAGCCGCTGCTCTATCTGCTTCTTGTGACAGCTGCTTCTTGTCT 60
QY 61 GTGAGAGGCGAGGGGTGGCAAGCTTTCAGGAATGATCCACAGAGGTATCCAGGAGCTT 120
DB 61 GTGAGAGGCGAGGGGTGGCAAGCTTTCAGGAATGATCCACAGAGGTATCCAGGAGCTT 120
QY 121 GGAGAGTACCCCGAGGCTCTCTTCTGAGAACCAACCAACCAATGAAACCGGGCGAGAAATGA 180
DB 121 GGAGAGTACCCCGAGGCTCTCTTCTGAGAACCAACCAACCAATGAAACCGGGCGAGAAATGA 180
QY 181 GCGAGAGCTCCCAACCATCTCTATGACGCAAAAGATGTTCAGATGACAGCTCCCGGAG 240
DB 181 GCGAGAGCTCCCAACCATCTCTATGACGCAAAAGATGTTCAGATGACAGCTCCCGGAG 240
QY 241 CTGACACTACAGCCGCTCTCTGACAGAGCGCCATGCGCAGCGCAAGCGGTCAACCGAG 300
DB 241 CTGACACTACAGCCGCTCTCTGACAGAGCGCCATGCGCAGCGCAAGCGGTCAACCGAG 300
QY 301 TTGAGTGTCTCCGCGCAAGTGGCGCCCGCGGCTGCTGCTCCCAACCGCATCGGCGGTG 360
DB 301 TTGAGTGTCTCCGCGCAAGTGGCGCCCGCGGCTGCTGCTCCCAACCGCATCGGCGGTG 360
QY 361 AAGTGTGTGGCGCCCGCAACCGCATTTTCCGCTGATCCCGATGCGTCAACCGCGGAG 420
DB 361 AAGTGTGTGGCGCCCGCAACCGCATTTTCCGCTGATCCCGATGCGTCAACCGCGGAG 420
QY 421 CGGAGTGAAGCTGTGTGCGCGCGGGCGCGCGCGCTGCGCAAGTGTGCTGTGAGT 480
DB 421 CGGAGTGAAGCTGTGTGCGCGCGGGCGCGCGCGCTGCGCAAGTGTGCTGTGAGT 480
QY 481 GCTGTGTGAAGTGAAGCGCTTCAACCGCTTCCCAACCAAGTGTGAGTCAAGACTTC 540
DB 481 GCTGTGTGAAGTGAAGCGCTTCAACCGCTTCCCAACCAAGTGTGAGTCAAGACTTC 540
QY 541 GGGCGGAGACCGCGCGCGCGCAAGAGGTGCAAGCGCGCGCGCGCGCGCGCGGAGCC 600
DB 541 GGGCGGAGACCGCGCGCGCGCAAGAGGTGCAAGCGCGCGCGCGCGCGCGCGGAGCC 600
QY 601 AAAGCCAAACGAGCGGAGCTGAGAAAGCTTACTAG 636
DB 601 AAAGCCAAACGAGCGGAGCTGAGAAAGCTTACTAG 636

RESULT 2
AF326740 638 bp mRNA linear ROD 28-FEB-2001
LOCUS Mus musculus sclerostin mRNA, complete cds.
DEFINITION AF326740
ACCESSION AF326740
VERSION AF326740.1 GI:13161022
KEYWORDS
SOURCE Mus musculus.
ORGANISM
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS
Brunkow, M.E., Gardner, J.C., Van Ness, J., Paepker, B.W.,
Kovacevich, B.R., Prohl, S., Skonier, J.E., Zhao, L., Sabo, P.J.,
Fu, Y.H., Altsch, R.S., Gillett, L., Colbert, T., Tacconi, P., Galas, D.,
Hamersma, H., Beighton, P. and Mulligan, J.T.

TITLE Bone dysplasia sclerosteosis results from loss of the SOST gene product, a novel cystine knot-containing protein
JOURNAL Am. J. Hum. Genet. 68 (3), 577-589 (2001)
MEDLINE 21090529
PUBMED 11179006
REFERENCE 2 (bases 1 to 638)
AUTHORS Brunkow, M.E., Gardner, J.C., Van Ness, J., Paepker, B.W.,
Kovacevich, B.R., Prohl, S., Skonier, J.E., Zhao, L., Sabo, P.J.,
Fu, Y.H., Altsch, R.S., Gillett, L., Colbert, T., Tacconi, P., Galas, D.,
Hamersma, H., Beighton, P. and Mulligan, J.T.

TITLE Direct Submission
JOURNAL Submitted (07-DEC-2000) Genomics, Celltech Chiroscience Inc., 1621
220th St. SE, Bothell, WA 98021, USA

FEATURES
source
1. .636
/organism="Mus musculus"
/strain="129SV/J"
/db_xref="taxon:10090"
/chromosome="11"

CDS

1. .636
/codon_start=1
/product="sclerostin"
/protein_id="AAK13455.1"
/db_xref="GI:13161023"
/translation="MOSSLAPCLICLVHAAFCVAGGQWAFRNDATVETPLGSEYP
EPPENNQTMRAENGRPHHPYDPAKGVSEVSCRELHYRPLTDGCRSAKVELV
CSGCGPRLPLPAIGHVKKWRPNRGPFRIPRYAQRVQLCPGGAAPRSKRALV
ASCKKRLTFHNSLEKDPFETPARQKRPKRPARKAKAQAELNNAY"

BASE COUNT 114 a 224 c 209 g 91 t
ORIGIN

Query Match 92.0%; Score 585; DB 10; Length 638;
Best Local Similarity 99.8%; Pred. No. 1.8e-279; Indels 0; Gaps 0;
Matches 635; Conservative 0; Mismatches 1;

QY 1 ATGAGAGCCCTCACTAGCCGCTGCTCTATCTGCTTCTTGTGACAGCTGCTTCTTGTCT 60
DB 1 ATGAGAGCCCTCACTAGCCGCTGCTCTATCTGCTTCTTGTGACAGCTGCTTCTTGTCT 60
QY 61 GTGAGAGGCGAGGGGTGGCAAGCTTTCAGGAATGATCCACAGAGGTATCCAGGAGCTT 120
DB 61 GTGAGAGGCGAGGGGTGGCAAGCTTTCAGGAATGATCCACAGAGGTATCCAGGAGCTT 120
QY 121 GGAGAGTACCCCGAGGCTCTCTTCTGAGAACCAACCAACCAATGAAACCGGGCGAGAAATGA 180
DB 121 GGAGAGTACCCCGAGGCTCTCTTCTGAGAACCAACCAACCAATGAAACCGGGCGAGAAATGA 180
QY 181 GCGAGAGCTCCCAACCATCTCTATGACGCAAAAGATGTTCAGATGACAGCTCCCGGAG 240
DB 181 GCGAGAGCTCCCAACCATCTCTATGACGCAAAAGATGTTCAGATGACAGCTCCCGGAG 240
QY 241 CTGACACTACAGCCGCTCTCTGACAGAGCGCCATGCGCAGCGCAAGCGGTCAACCGAG 300
DB 241 CTGACACTACAGCCGCTCTCTGACAGAGCGCCATGCGCAGCGCAAGCGGTCAACCGAG 300
QY 301 TTGAGTGTCTCCGCGCAAGTGGCGCCCGCGGCTGCTGCTCCCAACCGCATCGGCGGTG 360
DB 301 TTGAGTGTCTCCGCGCAAGTGGCGCCCGCGGCTGCTGCTCCCAACCGCATCGGCGGTG 360
QY 361 AAGTGTGTGGCGCCCGCAACCGCATTTTCCGCTGATCCCGATGCGTCAACCGCGGAG 420
DB 361 AAGTGTGTGGCGCCCGCAACCGCATTTTCCGCTGATCCCGATGCGTCAACCGCGGAG 420
QY 421 CGGAGTGAAGCTGTGTGCGCGGGCGCGCGCGCTGCGCAAGTGTGCTGTGAGT 480
DB 421 CGGAGTGAAGCTGTGTGCGCGGGCGCGCGCGCTGCGCAAGTGTGCTGTGAGT 480
QY 481 GCTGTGTGAAGTGAAGCGCTTCAACCGCTTCCCAACCAAGTGTGAGTCAAGACTTC 540
DB 481 GCTGTGTGAAGTGAAGCGCTTCAACCGCTTCCCAACCAAGTGTGAGTCAAGACTTC 540
QY 541 GGGCGGAGACCGCGCGCGCGCAAGAGGTGCAAGCGCGCGCGCGCGCGCGGAGCC 600
DB 541 GGGCGGAGACCGCGCGCGCGCAAGAGGTGCAAGCGCGCGCGCGCGCGCGGAGCC 600

Db 541 GGGCCGAGACCGCGCGCCGAGAGGGTTCGACAGCCCGCGCCGCCCGGAGACC 600
Qy 601 AAGCCCAACGAGCGAGCTGAGAGCCCTACTAG 636
Db 601 AAGCCCAACGAGCGAGCTGAGAGCCCTACTAG 636

RESULT 3
AF326737
LOCUS
DEFINITION
Mus musculus sclerostin (Sost) gene, complete cds; and Meox1 gene,
partial sequence.
AF326737.2 GI:20150447

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Mus musculus.
Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathia; Muridae; Murinae; Mus.
1 (bases 1 to 81806)
Brunkow,M.E., Gardner,J.C., Van Ness,J., Paepert,B.W.,
Kovacevich,B.R., Prolli,S., Skonier,J.E., Zhao,L., Sabo,P.J.,
Fu,Y.H., Altsch,R.S., Gillett,L., Colbert,T., Tacconi,P., Galas,D.,
Hamerma,H., Beighton,P. and Mulligan,J.T.
Bone dysplasia sclerosteosis results from loss of the SOST gene
product, a novel cysteine knot-containing protein
Am. J. Hum. Genet. 68 (3), 577-589 (2001)
21090529
11179006

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
2 (bases 1 to 81806)
Staebling-Hampton,K., Prolli,S., Paepert,B.W., Zhao,L., Charmley,P.,
Brown,A., Gardner,J.C., Galas,D., Schatzman,R.C., Beighton,P.,
Papapoulos,S., Hamersma,H. and Brunkow,M.E.
A 52 kb deletion in the SOST - MEOX1 intergenic region on 17q12-q21
is associated with van Buchem disease in the Dutch population
Unpublished
3 (bases 1 to 17423)
Brunkow,M.E., Gardner,J.C., Van Ness,J., Paepert,B.W.,
Kovacevich,B.R., Prolli,S., Skonier,J.E., Zhao,L., Sabo,P.J.,
Fu,Y.H., Altsch,R.S., Gillett,L., Colbert,T., Tacconi,P., Galas,D.,
Hamerma,H., Beighton,P. and Mulligan,J.T.
Direct Submission
Submitted (07-DEC-2000) Genomics, Celltech Chiroscience Inc., 1631
220th St. SE, Bothell, WA 98021, USA
4 (bases 1 to 81806)
Staebling-Hampton,K., Prolli,S., Paepert,B.W., Zhao,L., Charmley,P.,
Brown,A., Gardner,J.C., Galas,D., Schatzman,R.C., Beighton,P.,
Papapoulos,S., Hamersma,H. and Brunkow,M.E.
Direct Submission
Submitted (02-AUG-2001) Genomics, Celltech R&D Inc., 1631 220th St
SE, Bothell, WA 98021, USA
Sequence update by submitter
On Apr 15, 2002 this sequence version replaced gi:13161013.
Location/Qualifiers
1. 81806
/organism="Mus musculus"
/strain="129SV/J"
/db_xref="taxon:10090"
/chromosome="11"
/map="between Sost and Meox1"
<8036..>11163
/gene="Sost"
join(<8036..8249,10742..>11163)
/gene="Sost"
/product="sclerostin"
join(8036..8249,10742..11163)
/gene="Sost"
/codon_start=1
/product="sclerostin"
/protein_id="AAK13452.1"
/db_xref="GI:13161014"
/translation="MPSIAPCLICLVHAAFCAGVGGMQAFNDATFVIGIGEP
EPPENQTMRAENGRRPHHPYDAKDVSEISCEMLITRTFLIDGPERSKAPVTELV

TITLE
JOURNAL
REMARK
COMMENT
FEATURES
source

gene
mrna
CDS

CSGCCPALLPNAIGRVKWRPNNGDPFRCIIDRYRORVOLLCPGAGAPRSRRVRLV
ASCKKRLTRFHNOSELKDPFETARPOKGRKPRGARGAANAELENAI"
81492..>81806
/gene="Meox1"
81492..>81806
/gene="Meox1"
/product="Meox1"

BASE COUNT 20982 a 20323 c 20010 g 20491 t
ORIGIN

Query Match 66.7%; Score 424; DB 10; Length 81806;
Best Local Similarity 100.0%; Pred. No. 8.9e-200;
Matches 424; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 213 AGATGTGTCGAGTACAGTGCAGGAGCTGACATACACCCCTTCTGACAGAGCC 272
Db 10740 AGATGTGTCGAGTACAGTGCAGGAGCTGACATACACCCCTTCTGACAGAGCC 10739

Qy 273 ATGCCGACAGCCGACAGCCGATGACAGGATTGTCCTCCGACAGTGCAGCCGCG 332
Db 10800 ATGCCGACAGCCGACAGCCGATGACAGGATTGTCCTCCGACAGTGCAGCCGCG 10859

Qy 333 GCTGCTGCCAAGCCGATGCGGCGGTGAAGTGTGCGGCCGGAACGAGCCGATTTCG 392
Db 10860 GCTGCTGCCAAGCCGATGCGGCGGTGAAGTGTGCGGCCGGAACGAGCCGATTTCG 10919

Qy 393 CTGCATCCCGGATGCTACCGCGGCGGAGCGGTGACGCTGCTGTCCCGGGGGCGCGC 452
Db 10920 CTGCATCCCGGATGCTACCGCGGCGGAGCGGTGACGCTGCTGTCCCGGGGGCGCGC 10979

Qy 453 GCGGCGCTGCGCAGAGTGCCTGTGAGCTCTGCAAGTGCAGAGCCGCTCACCCGCTT 512
Db 10980 GCGGCGCTGCGCAGAGTGCCTGTGAGCTCTGCAAGTGCAGAGCCGCTCACCCGCTT 11039

Qy 513 CCACACCAAGTGTGAGCTTCAGACTTCGCGCGGAGACCGCGCGCGCGCAGAAAGGTG 572
Db 11040 CCACACCAAGTGTGAGCTTCAGACTTCGCGCGGAGACCGCGCGCGCGCAGAAAGGTG 11099

Qy 573 CAAGCG 632
Db 11100 CAAGCG 11159

Qy 633 CTAG 636
Db 11160 CTAG 11163

RESULT 4
AC068782.2
WPCOMMENT
Sequence split into 5 fragments LOCUS AC068782 Accession AC068782
Fragment Name Begin End
AC068782_0 1 110000
AC068782_1 100001 210000
AC068782_2 200001 310000
AC068782_3 300001 410000
AC068782_4 400001 441130
Continuation (3 of 5) of AC068782 from base 200001 (AC068782 Mus musculus chromosome 11 c

Query Match 66.7%; Score 424; DB 2; Length 110000;
Best Local Similarity 100.0%; Pred. No. 8.5e-200;
Matches 424; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 213 AGATGTGTCGAGTACAGTGCAGGAGCTGACATACACCCCTTCTGACAGAGCC 272
Db 5660 AGATGTGTCGAGTACAGTGCAGGAGCTGACATACACCCCTTCTGACAGAGCC 56659

Qy 273 ATGCCGACAGCCGACAGCCGATGACAGGATTGTCCTCCGACAGTGCAGCCGCG 332
Db 5660 ATGCCGACAGCCGACAGCCGATGACAGGATTGTCCTCCGACAGTGCAGCCGCG 56719

Qy 333 GCTGCTGCCAAGCCGATGCGGCGGTGAAGTGTGCGGCCGGAACGAGCCGATTTCG 392

Db	Accession	Source	Organism	Reference	Title	Journal	Authors
Db	56720	GI50760	CCAAACGCGCATGAGGCGGTAAAGTGTGGCGCCGGAACGGAACGGATTTCCG	56779			
Qy	393	CTGCATCCCGGATCGCTACCCGCGCGAGCGGGTGTCACTGCTGTGTCCCGCGGGCGGGCGGC	452				
Db	56780	CTGCATCCCGGATCGCTACCCGCGCGAGCGGGTGTCACTGCTGTGTCCCGCGGGCGGGCGGC	56839				
Qy	453	GCCGCGCTCGCCCAAGTGTGTCTGTGGCGCTGTGCAAGTGCAGCGCTTACCCGCTT	512				
Db	56840	GCCGCGCTCGCCCAAGTGTGTCTGTGGCGCTGTGCAAGTGCAGCGCTTACCCGCTT	56899				
Qy	513	CCACAACCACTGCGAGCTCAAGGACTTTGGGCGCGGAGACCGCGCGCGCAGAAAGGCTG	572				
Db	56900	CCACAACCACTGCGAGCTCAAGGACTTTGGGCGCGGAGACCGCGCGCGCAGAAAGGCTG	56959				
Qy	573	CAAGCTA	632				
Db	56960	CAAGCTA	57019				
Qy	633	CTAG 636					
Db	57020	CTAG 57023					
RESULT 5							
LOCUS	AC012296	205277 bp	DNA	linear	HTG 16-AUG-2002		
DEFINITION	Mus musculus chromosome 11 clone RP23-346P7 map 11, *** SEQUENCING						
ACCESSION	AC012296	10 GI:2267737					
VERSION	AC012296.10						
KEYWORDS	HTG, HTGS, PHASE1, HTGS_FUZZTOP, HTGS_ACTIVIFIN.						
SOURCE	house mouse.						
ORGANISM	Mus musculus						
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.						
AUTHORS	1 (bases 1 to 205277)						
TITLE	Birren, B., Nusbaum, C. and Lander, E.						
JOURNAL	Mus musculus chromosome 11, clone RP23-346P7						
REFERENCE	Unpublished						
AUTHORS	2 (bases 1 to 205277)						
TITLE	Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M.,						
JOURNAL	Balwin, J., Barna, N., Beckwith, R., Boguslavsky, L., Bonkshalev, B.,						
AUTHORS	Brown, A., Castle, A., Colangelo, M., Collins, S., Collymore, A.,						
TITLE	Cooke, P., DeBelland, K., Dewar, K., Domino, M., Donnell, L., Doyle, M.,						
JOURNAL	Ferreira, P., Fitzhugh, W., Forrest, C., Funke, R., Gage, D., Horton, L.,						
AUTHORS	Howland, J., Gadyana, S., Grant, G., Hagos, B., Hearford, A., Klein, J.,						
TITLE	Lehoczky, J., Lien, C., Locks, K., MacDonald, P., Marquis, N.,						
AUTHORS	McEwan, P., McGuire, A., McKernan, K., McLaughlin, J., Meldrum, J.,						
TITLE	Morrow, J., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P.,						
AUTHORS	Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P.,						
TITLE	Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talmas, J.,						
AUTHORS	Testa, S., Turrell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X.,						
TITLE	Wymann, D., Ye, W. J., Zimmer, A. and Zody, M.						
JOURNAL	Submitted (22-OCT-1999) Whitehead Institute/MIT Center for Genome						
AUTHORS	Research, 320 Charles Street, Cambridge, MA 02141, USA						
REFERENCE	3 (bases 1 to 205277)						
AUTHORS	Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,						
TITLE	Barna, N., Bastien, V., Bloom, F., Boguslavsky, L., Bonkshalev, B.,						
JOURNAL	Camrath, J., Chang, J., Chazaro, F., Choepel, J., Collymore, A.,						
AUTHORS	Cook, A., Cooke, P., DeBelland, K., Dewar, K., Diaz, J. S., Dodge, S.,						
TITLE	Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J.,						
AUTHORS	Gadyana, S., Gadyana, S., Grant, G., Hagos, B., Hearford, A., Klein, J.,						
TITLE	Horton, L., Hu, M., Iliev, I., Johnson, R., Jones, C., Kamat, A.,						
AUTHORS	Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K.,						
TITLE	Liu, G., Maclean, C., MacDonald, P., Major, J., Matthews, C.,						
AUTHORS	McCarthy, M., Meldrum, J., Menus, L., Mhova, T., Mlewa, V.,						
TITLE	Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H.,						
AUTHORS	O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,						

TITLE
JOURNAL
COMMENT

Testes, S., Theodore, J., Tophan, K., Travers, M., Vassiliev, H.,
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J.,
 Zembek, I., Zimmer, A. and Zody, M.
 Direct Submission
 Submitted (16-AUG-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Aug 16, 2002 this sequence version replaced gi:20177756.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RW/RepeatMasker.html>
 ----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: MIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information -----

Center project name: L3205
Center clone name: 346_P_7

* NOTE: This is a "working draft" sequence. It currently
* consists of 10 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the configs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 92178: contig of 92178 bp in length
 9219 92278: gap of 100 bp
 92279 128704: contig of 36426 bp in length
 128705 128804: gap of 100 bp
 128805 140272: contig of 11468 bp in length
 140273 140373: gap of 100 bp
 140373 147001: contig of 6629 bp in length
 147002 147101: gap of 100 bp
 147102 170176: contig of 23075 bp in length
 170177 170276: gap of 100 bp
 170277 172609: contig of 333 bp in length
 173610 173709: gap of 100 bp
 173710 175722: contig of 2013 bp in length
 175723 175822: gap of 100 bp
 175823 181949: contig of 6127 bp in length
 181950 182049: gap of 100 bp
 182050 191157: contig of 9108 bp in length
 191258 191257: gap of 100 bp
 191258 205277: contig of 14020 bp in length

FEATURES
SOURCE

```

/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="1"
/map="11"
/clone="RP23-346B7"
/clone_1lb="RBC1-23
Female Mouse BAC"
53425 a 49464 c 50390 g 50980 t 1018 others
BASE COUNT
ORIGIN

```

Query Match	66.7%;	Score 424;	DB 2;	Length 205277;
Best Local Similarity	100.0%;	Pred. No. 7.8e-200;		

Matches	424;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
213	AGATGTGTCGAGTACACTGCGGAGCTGCATACACCCGCTTCTCTGACAGCGGCC	212							

Db 57866 AGATGTCTCCGATACAGCTGCGCGGAGCTGCACTACACCCGGCTTCTCTGACAGACGCGCC 57925

273 ATCCCGAGCGCCAAAGCCGCTACCCAGATTGGTGTCTCCGGCAGTCCGGCCCGCGCG 332
QY
57926 ATCCCGAGCGCCAAAGCCGCTACCCAGATTGGTGTCTCCGGCAGTCCGGCCCGCGCG 57985
nb

333 GCTGCTGCCAAGCCATCGGGCGCTGAAGTGTGCGGCCGAA CGGACCGGATTTCCG 392

Db 57986 GCTGCTGCCCAACGCCCATCGGCGCGGTGAAGTGTGGCGCCCGAAGGACCGGATTCCG 58045


```

repeat_region      /rpt_family="(CA)n"
                    15971..16099
repeat_region      /rpt_family="L2"
                    16336..16404
repeat_region      /rpt_family="BC1_MM"
                    16419..16550
repeat_region      /rpt_family="B1-F"
                    17412..17518
repeat_region      /rpt_family="B1-F"
                    17538..17710
repeat_region      /rpt_family="B3A"
                    18267..18397
repeat_region      /rpt_family="B3A"
                    18508..18679
repeat_region      /rpt_family="B3"
                    18939..18982
repeat_region      /rpt_family="B3"
                    18939..18982
repeat_region      /rpt_family="(TTPAG)n"
                    20446..20476
repeat_region      /rpt_family="(TG)n"
                    20446..20476
repeat_region      /rpt_family="B4A"
                    20593..20889
repeat_region      /rpt_family="B4A"
                    20593..20889
repeat_region      /rpt_family="PB1D10"
                    22076..22187
STS
repeat_region      /standard_name="163518"
                    22076..22187
repeat_region      /rpt_family="MIR"
                    22748..22899
repeat_region      /rpt_family="B4A"
                    23282..23305
repeat_region      /rpt_family="(A)n"
                    23282..23305
repeat_region      /rpt_family="B2"
                    23322..23355
repeat_region      /rpt_family="B2"
                    23322..23355
repeat_region      /rpt_family="B2"
                    23381..23525
repeat_region      /rpt_family="B2"
                    24185..24255
repeat_region      /rpt_family="(TG)n"
                    24921..25202
repeat_region      /rpt_family="Lx6"
                    25271..25346
repeat_region      /rpt_family="B3"
                    25352..25444
repeat_region      /rpt_family="PB1D10"
                    26173..26320
repeat_region      /rpt_family="L1MD"
                    26434..26601
repeat_region      /rpt_family="L1MD"
                    26720..26801
repeat_region      /rpt_family="RSINE1"
                    26802..26906
repeat_region      /rpt_family="B4A"
                    27067..27454
repeat_region      /rpt_family="PB1D7"
                    27555..27575
repeat_region      /rpt_family="L1_MM"
                    27956..28217
repeat_region      /rpt_family="B4"
                    28275..28459
repeat_region      /rpt_family="L1MD"
                    28460..28615
repeat_region      /rpt_family="B3A"
                    28837..29077
repeat_region      /rpt_family="B2"
                    29241..29376
repeat_region      /rpt_family="B4A"
                    29323..29410
repeat_region      /rpt_family="PB1D10"

```

```

repeat_region      29438..29691
repeat_region      /rpt_family="B4A"
                    29696..29823
repeat_region      /rpt_family="(CA)n"

Query Match
Best Local Similarity 100.0%; Pred. No. 7,8e-200;
Matches 424; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 213 AGATGTGTCGAGTACAGCTGTCGCGGAGCTGCACTACACCCGCTTCTGTACACAGCGCC 272
Db 57847 AGATGTGTCGAGTACAGCTGTCGCGGAGCTGCACTACACCCGCTTCTGTACACAGCGCC 57906
QY 273 ATGCCGACGCGCCAGCCGCGGACGAGTTGTGTGCTCCGCGGACGAGCGCCGCGC 332
Db 57907 ATGCCGACGCGCCAGCCGCGGACGAGTTGTGTGCTCCGCGGACGAGCGCCGCGC 57966
QY 333 GCTGCTGCCCAACGCGCATCGGCGCGGTGAAGTGTGCGCGCCGAAACGAGCTTCCG 392
Db 57967 GCTGCTGCCCAACGCGCATCGGCGCGGTGAAGTGTGCGCGCCGAAACGAGCTTCCG 58026
QY 339 CTGCATCCCGGATGCTACCGCGCGGACGCGGTGCTGTGCTTCCCGGGCGCGGC 452
Db 58027 CTGCATCCCGGATGCTACCGCGCGGACGCGGTGCTGTGCTTCCCGGGCGCGGC 58086
QY 453 GCGCGCTCGCGCAAGGCGGTGCTGTGCTGCTGTGCAAGTGCAGAGCGCTCACCCGCTT 512
Db 58087 GCGCGCTCGCGCAAGGCGGTGCTGTGCTGCTGTGCAAGTGCAGAGCGCTCACCCGCTT 58146
QY 513 CCACAACGATCGAGTCAAGGCTCAAGGCTCGGAGACCGCGGCGCGCAAGAGGCTCG 572
Db 58147 CCACAACGATCGAGTCAAGGCTCAAGGCTCGGAGACCGCGGCGCGCAAGAGGCTCG 58206
QY 573 CAAGCGCGCGCGCGCGCGCGCGGAGCCAAAGCCAGCGGAGTGTGAAGAGCGCTA 632
Db 58207 CAAGCGCGCGCGCGCGCGCGCGGAGCCAAAGCCAGCGGAGTGTGAAGAGCGCTA 58266
QY 633 CTAG 636
Db 58267 CTAG 58270

RESULT 7
AL591145
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

```

Mouse DNA sequence from clone Rp23-3987 on chromosome 11, complete sequence.
 AL591145
 AL591145.24 GI:22474402
 HTG.
 house mouse.
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 198508)
 Garner, P.
 Direct Submission
 Submitted (22-AUG-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
 On Aug 23, 2002 this sequence version replaced gi:2203099.
 ----- Genome Center
 Center: Wellcome Trust Sanger Institute
 Center code: SC
 Web site: <http://www.sanger.ac.uk>
 Contact: humquerry@sanger.ac.uk

 During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WormPEP; Information on the WormPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-398F7 is from the RP23 Mouse PAC Library constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm> VECTOR: pBAC3.6.

FEATURES

source

Location/Qualifiers
1..198508 "Mus musculus"
/db_xref="taxon:10090"
/chromosome="11"
/clone="RP23-398F7"
/clone_11b="RP23-23"

BASE COUNT 48902 a 47818 c 50456 g 51332 t
ORIGIN

Query Match 58.6%; Score 373; DB 10; Length 198508;
Best Local Similarity 99.8%; Pred. No. 1.7e-174;
Matches 423; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 213 AGATGTGTCCGAGTACAGCTGCGCCGAGCTGCACTACACCCGCTTCTTACAGAGCGGCC 272
Db 187744 AGATGTGTCCGAGTACAGCTGCGCCGAGCTGCACTACACCCGCTTCTTACAGAGCGGCC 187803
QY 273 ATCCCGCAGCGCCAGCGGCTGCAACGAGTGTGTGCTCCGAGCGGCGCCCGCGCG 332
Db 187804 ATCCCGCAGCGCCAGCGGCTGCAACGAGTGTGTGCTCCGAGCGGCGCCCGCGCG 187863
QY 333 GCTGTGCTCCCAAGCGCCATCGGCGCGGTGAAGTGTGTGCGCCGCAACGAGCGGATTTCCG 392
Db 187864 GCTGTGCTCCCAAGCGCCATCGGCGCGGTGAAGTGTGTGCGCCGCAACGAGCGGATTTCCG 187923
QY 393 CTGCATCCCGGAGTCTTACCGCGCGGCGGAGCTGTCTGTGCGCGGCGGCGCGCG 452
Db 187924 CTGCATCCCGGAGTCTTACCGCGCGGCGGAGCTGTCTGTGCGCGGCGGCGCGCG 187983
QY 453 GCGCGCTCGCGCAGGAGTGTCTGTGCGCTGTGCAAGTGTGCAAGCGGCTTCAACCGCTT 512
Db 187984 GCGCGCTCGCGCAGGAGTGTCTGTGCGCTGTGCAAGTGTGCAAGCGGCTTCAACCGCTT 188043
QY 513 CCACACCAAGTGTGAGCTTCTGCGCGGAGACCGCGCGCGCGCGCGAGAGGCTCG 572
Db 188044 CCACACCAAGTGTGAGCTTCTGCGCGGAGACCGCGCGCGCGCGCGAGAGGCTCG 188103
QY 573 CAGCG 632
Db 188104 CAGCG 188163
QY 633 CTAG 636
Db 188164 CTAG 188167

RESULT 8
AC023810/c 51575 bp DNA linear HTG 04-NOV-2000
LOCUS Mus musculus chromosome 11 clone RP23-252B10, *** SEQUENCING IN
DEFINITION PROGRESS *** 35 unordered pieces.
AC023810
AC023810
AC023810.6 GI:11079371
KEYWORDS HTG; HTGS; PHASE1.
SOURCE Mus musculus.
ORGANISM Mus musculus

REFERENCE

AUTHORS

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 51575)
Metzker, D.L., Lewis, L.R., Hume, J., Edwards, C., Harris, C., Dederich, D., Thomas, S., Okwuonu, G., Carlack, C., Garner, T., Addison, S., Pace, A., Williams, G., Bonnin, D., Brooks, A., Brown, J., Bunay, C., Bunac, C., Burkett, C., Chacko, J., Chen, G., Chen, Z., Cox, C., Davis, C., Delgado, O., Ding, Y., Dugan-Rocha, S., Fernandez, C., Ferraguto, D., Forcum-Taney, J., Gill, R., Gorenell, J.H., Gunaratne, P., Haller, G., Hernandez, J., Hogues, M., Hosak, H., Hou, X., Huber, J., Jackson, L., Jia, Y., Kelly, J., Kelly, S., Kovar, C., Liu, J., Liu, W., Louised, H., Lozano, R.J., Martin, R., Massey, E., McLeod, M.P., Mei, G., Moore, S., Morgan, M., Morris, S., Neal, D., Nelson, A., Nguyen, R., Nguyen, N., Ogih, M., Parish, B., Perez, L., Reiter, D., Say, J., Shen, H., Vasquez, L., Watlington, S., Williamson, A., Wrensford, G., Zhou, X., Bouck, J., Hodgson, A., Murry, D.M., Rives, M., Scherer, S., Sodergren, E., Weinstock, G., Worley, K. and Gibbs, R.
Direct Submission
2 (bases 1 to 51575)
Worley, K.C.
Submitted (18-FEB-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Nov 3, 2000 this sequence version replaced gi:8248603.
----- Genome Center -----
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information -----
Center project name: MABO
Center clone name: RP23-252B10
----- Summary Statistics -----
Sequencing vector: M13; L08821
Chemistry: Dye-terminator Big Dye: 76% of reads
Chemistry: Dye-terminator Big Dye: 24% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 42708 bases at least Q40
Consensus quality: 59319 bases at least Q30
Consensus quality: 66755 bases at least Q20
Estimated insert size: 39143; sum-of-contigs estimation
Quality coverage: 0x in Q20 bases; agorose-fp estimation
Quality coverage: 0.4x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbankdraft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 35 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
2683: contig of 2682 bp in length
2783: gap of unknown length
2783: contig of 3111 bp in length
5893: gap of unknown length
5893: contig of 1446 bp in length
7439: gap of unknown length
7439: contig of 1767 bp in length
7540: gap of unknown length
9306: contig of 1767 bp in length
9307: gap of unknown length
9405: contig of 1856 bp in length
11263: gap of unknown length
11263: contig of 1374 bp in length
11363: gap of unknown length
12736: contig of 1116 bp in length
12837: gap of unknown length
13952: contig of 1154 bp in length
13953: gap of unknown length
14053: contig of 1154 bp in length
15207: gap of unknown length
15306: gap of unknown length

```

* 15307 16973: contig of 1667 bp in length
* 16974 17073: gap of unknown length
* 17074 18081: contig of 1008 bp in length
* 18082 18181: gap of unknown length
* 18182 19630: contig of 1449 bp in length
* 19631 19730: gap of unknown length
* 19731 20732: contig of 1002 bp in length
* 20733 20832: gap of unknown length
* 20833 22555: contig of 1723 bp in length
* 22556 22656: gap of unknown length
* 22656 23846: contig of 1191 bp in length
* 23847 23946: gap of unknown length
* 23947 25489: contig of 1543 bp in length
* 25490 25590: gap of unknown length
* 25590 26913: contig of 1324 bp in length
* 26914 27013: gap of unknown length
* 27014 28158: contig of 1145 bp in length
* 28159 28258: gap of unknown length
* 28259 29424: contig of 1166 bp in length
* 29425 29525: gap of unknown length
* 29525 30594: contig of 1070 bp in length
* 30595 30694: gap of unknown length
* 30695 32013: contig of 1319 bp in length
* 32014 32114: gap of unknown length
* 32114 33542: contig of 1429 bp in length
* 33543 33642: gap of unknown length
* 33643 34705: contig of 1063 bp in length
* 34706 34805: gap of unknown length
* 34806 35911: contig of 1106 bp in length
* 35912 36011: gap of unknown length
* 36012 37239: contig of 1228 bp in length
* 37240 37339: gap of unknown length
* 37340 38613: contig of 1274 bp in length
* 38614 38713: gap of unknown length
* 38714 39885: contig of 1172 bp in length
* 39886 39985: gap of unknown length
* 39986 41336: contig of 1351 bp in length
* 41337 41436: gap of unknown length
* 41437 42561: contig of 1125 bp in length
* 42562 42661: gap of unknown length
* 42662 43685: contig of 1024 bp in length
* 43686 43785: gap of unknown length
* 43786 45017: contig of 1232 bp in length
* 45018 45117: gap of unknown length
* 45118 46507: contig of 1390 bp in length
* 46508 46607: gap of unknown length
* 46608 47781: contig of 1174 bp in length
* 47782 47881: gap of unknown length
* 47882 49024: contig of 1143 bp in length
* 49025 49124: gap of unknown length
* 49125 50392: contig of 1268 bp in length
* 50393 50492: gap of unknown length
* 50493 51575: contig of 1083 bp in length.

```

FEATURES
source
1. .51575
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="11"
/clone="RP23-252B10"

BASE COUNT 12565 a 11207 c 11347 g 13002 t 3454 others
ORIGIN

Query Match 33.6%: Score 214; DB 2; Length 51575;
Best Local Similarity 100.0%; Pred. No. 2.1e-95;
Matches 214; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 ATGAGCCCTCCTAGCCCGCTGCTCATCTGCTTGTGACGCTGCTTGTGCT 60
DB 8375 ATGAGCCCTCCTAGCCCGCTGCTCATCTGCTTGTGACGCTGCTTGTGCT 8316
QY 61 GTGAGGCGCCAGGGGTGGCAAGCTTCAGAGATATGTCACAGAGGTTCATCCAGGCTT 120
DB 8315 GTGAGGCGCCAGGGGTGGCAAGCTTCAGAGATATGTCACAGAGGTTCATCCAGGCTT 8256

```

```

QY 121 GGAGAGTACCCCGAGCCTCTCTCTGAGAGCAACGAGCCATGAACCGGGGCGAGAAATGA 180
DB 8255 GGAGAGTACCCCGAGCCTCTCTCTGAGAGCAACGAGCCATGAACCGGGGCGAGAAATGA 8196
QY 181 GGCAGACCTCCCGACCAATCCCTATGAGCCCAAG 214
DB 8195 GGCAGACCTCCCGACCAATCCCTATGAGCCCAAG 8162

```

RESULT 9
AF326741 674 bp mRNA linear ROD 28-FEB-2001
LOCUS
DEFINITION Rattus norvegicus sclerostin mRNA, complete cds.
ACCESSION AF326741
VERSION AF326741.1 GI:13161025
KEYWORDS
SOURCE Rattus norvegicus.
ORGANISM Rattus norvegicus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE
AUTHORS 1 (bases 1 to 674)
Brunkow, M.E., Gardner, J.C., Van Nees, J., Paepker, B.W.,
Kovacevich, B.R., Prohl, S., Skonier, J.E., Zhao, L., Sabo, P.J.,
Fu, Y.H., Altsch, R.S., Gillett, L., Colbert, T., Tacconi, P., Galas, D.,
Hamersma, H., Beighon, P. and Mulligan, J.T.
Bone dysplasia sclerosteosis results from loss of the SOST gene
product, a novel cystine knot-containing protein
Am. J. Hum. Genet. 68 (3), 577-589 (2001)

TITLE
JOURNAL MEDLINE 21090529
PUBMED 1179006
2 (bases 1 to 674)
Brunkow, M.E., Gardner, J.C., Van Nees, J., Paepker, B.W.,
Kovacevich, B.R., Prohl, S., Skonier, J.E., Zhao, L., Sabo, P.J.,
Fu, Y.H., Altsch, R.S., Gillett, L., Colbert, T., Tacconi, P., Galas, D.,
Hamersma, H., Beighon, P. and Mulligan, J.T.
Direct Submission
Submitted (07-DEC-2000) Genomics, Celltech Chiroscience Inc., 1631
220th St. SE, Bothell, WA 98021, USA

JOURNAL
TITLE
JOURNAL
220th St. SE, Bothell, WA 98021, USA

FEATURES
source
1. .674
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
1. .32
33. .674
/codon_start=1
/product="sclerostin"
/protein_id="AAK13456.1"
/db_xref="GI:13161026"
/translation="MOELAPLCLLVHAFVAVESQGMQAFKNDATETIIPGLREY
EPPELENNOTNRAENGSRPHHPYDTHVVSQRELAHYRFPVDDGPRSAKPYR
LVCSGQCPARLPLNATIGRVKWRPNQSPDRCLPDYRRAORVOLLPGGAAPRSKVR
LVASCCKKRLTRFNOSLDFSPETARPKGRPRPRRAGAAANOAELENNAY"

BASE COUNT 128 a 240 c 208 g 98 t
ORIGIN

Query Match 16.4%: Score 104; DB 10; Length 674;
Best Local Similarity 99.4%; Pred. No. 1.8e-40;
Matches 154; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 388 TTCGGTGGATCCCGATCGGTACCGCGCGAGGCGAGTGTGTGCGCCGCGGCGC 447
DB 426 TTCGGTGGATCCCGATCGGTACCGCGCGAGGCGAGTGTGTGCGCCGCGGCGC 485
QY 448 GCGGCGCGCGGTCCGCGAGGTGCTGTGAGCTGTGCAAGTCAAGCGCTTACC 507
DB 486 GCGGCGCGCGGTCCGCGAGGTGCTGTGAGCTGTGCAAGTCAAGCGCTTACC 545
QY 508 CGCTTCACACCAATCGGAGTCAAGACTTGG 542
DB 546 CGCTTCACACCAATCGGAGTCAAGACTTGG 580

```

Db 99 PVTIVSGGCGPARLLPNAIGRGKWRPSPGDPFRCTIDRYRAORVOLLCPGGEAPAPARK 158
 QY 136 VRLVASCCKRLLTFPHNOSELKDFTEAARPOKGRKPRPARSAKANCAELENNAY 190
 Db 159 VRLVASCCKRLLTFPHNOSELKDFTEAARPOKGRKPRPARSAKANCAELENNAY 213

RESULT 10

AAM73067 standard; Protein; 139 AA.

AAM73067;

06-NOV-2001 (first entry)

Human bone marrow expressed probe encoded protein SEQ ID NO: 33373.

Human; bone marrow expressed exon; gene expression analysis; probe; microarray; cancer; leukaemia; lymphoma; myeloma.

Homo sapiens.

WO200157276-A2.

09-AUG-2001.

30-JAN-2001; 2001WO-US00668.

04-FEB-2000; 2000US-0180312.

26-MAY-2000; 2000US-0207456.

30-JUN-2000; 2000US-0608408.

03-AUG-2000; 2000US-0632366.

21-SEP-2000; 2000US-0234687.

27-SEP-2000; 2000US-0236359.

04-OCT-2000; 2000GB-0024263.

(MOLE-) MOLECULAR DYNAMICS INC.

Penn SG, Hanzel DK, Chen W, Rank DR;

WPI; 2001-488900/53.

Human genome-derived single exon nucleic acid probes useful for

analyzing gene expression in human bone marrow -

Example 4; SEQ ID NO: 33373; 658bp + Sequence Listing; English.

The present invention provides a number of single exon nucleic acid

probes which are derived from genomic sequences expressed in the human

bone marrow. They can be used to measure gene expression in bone marrow

samples, which may enable the improved diagnosis and treatment of cancers

such as lymphoma, leukaemia and myeloma. The present sequence is a

protein encoded by one of the probes of the invention.

Sequence 139 AA;

Query Match 73.2%; Score 139; DB 22; Length 139;

Best Local Similarity 100.0%; Pred. No. 3.4e-137;

Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 52 VSEYSGRELFHTYVTDPGRSAKPYTELVCSCGCCPARLLPNAIGRGKWRPSPGDPFRCT 111

Db 1 VSEYSGRELFHTYVTDPGRSAKPYTELVCSCGCCPARLLPNAIGRGKWRPSPGDPFRCT 60

QY 112 IPDRYRAORVOLLCPGGEAPAPARKVLAASCKCKRLTFPHNOSELKDFTEAARPOKGRK 171

Db 61 IPDRYRAORVOLLCPGGEAPAPARKVLAASCKCKRLTFPHNOSELKDFTEAARPOKGRK 120

QY 172 PRRPARSAKANCAELENNAY 190

Db 121 PRRPARSAKANCAELENNAY 139

RESULT 11

ABG42911 standard; Peptide; 139 AA.

ABG42911;

19-AUG-2002 (first entry)

Human peptide encoded by genome-derived single exon probe SEQ ID 32576.

Human; single exon probe; asthma; lung cancer; COPD; ILD;

chronic obstructive pulmonary disease; interstitial lung disease;

familial idiopathic pulmonary fibrosis; neurofibromatosis;

tuberosus sclerosis; Gaucher's disease; Niemann-Pick disease;

Hernandez-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;

pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;

pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;

primary ciliary dyskinesia; pulmonary hypertension;

hyaline membrane disease.

Homo sapiens.

WO200186003-A2.

15-NOV-2001.

30-JAN-2001; 2001WO-US00665.

04-FEB-2000; 2000US-180312P.

26-MAY-2000; 2000US-207456P.

30-JUN-2000; 2000US-0608408.

03-AUG-2000; 2000US-0632366.

21-SEP-2000; 2000US-234687P.

27-SEP-2000; 2000US-236359P.

04-OCT-2000; 2000GB-0024263.

(MOLE-) MOLECULAR DYNAMICS INC.

Penn SG, Hanzel DK, Chen W, Rank DR;

WPI; 2002-114183/15.

Spatially-addressable set of single exon nucleic acid probes, used to

measure gene expression in human lung samples -

Claim 27; SEQ ID No 32576; 634bp; English.

The invention relates to a spatially-addressable set of single exon

nucleic acid probes for measuring gene expression in a sample derived

from human lung comprising single exon nucleic acid probes having one of

12614 nucleic acid sequences mentioned in the specification, or their

complements or the 12387 open reading frames derived from the 12614

probes. Also included are a microarray comprising the novel set of

probes; the novel set of probes which hybridise at high stringency to a

nucleic acid expressed in the human lung; measuring gene expression in a

sample derived from human lung, comprising (a) contacting the array with

a collection of detectably labeled nucleic acids derived from human lung

mRNA, and (b) measuring the label detectably bound to each probe of

the array; identifying exons in a eukaryotic genome, comprising

(a) algorithmically predicting at least one exon from genomic sequences

of the eukaryote; and (b) detecting specific hybridisation of detectably

labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,

having a fragment identical to the predicted exon, the probe is included

in the above mentioned microarray; assigning exons to a single gene,

comprising (a) identifying exons from genomic sequence by the method

above and (b) measuring the expression of each of the exons in several

tissues and/or cell types using hybridisation to a single exon

microarrays having a probe with the exon, where a common pattern of

expression of the exons in the tissues and/or cell types indicates that

the exons should be assigned to a single gene; a peptide comprising one

of 12011 sequences, mentioned in the specification, or encoded by the

probes/open reading frames (ORF). The probes are used for gene

expression analysis, and for identifying exons in a gene, particularly

CC using human lung derived mRNA and for the study of lung diseases
 CC such as asthma, lung cancer, chronic obstructive pulmonary disease
 CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
 CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
 CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
 CC haemangiosarcoma, pulmonary histiocytosis, lymphangioleiomyomatosis,
 CC pulmonary alveolar proteinosis, Karagen syndrome, fibrocystic
 CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
 CC and hyaline membrane disease. The present sequence is a peptide/protein
 CC encoded by a single exon probe of the invention.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 139 AA;

Query Match 73.2%; Score 139; DB 23; Length 139;
 Best Local Similarity 100.0%; Pred. No. 3.4e-137;
 Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 52 VSEYSCRELHFTRYVTDGPGCRSAKPVTELVCSGCGPARLLPNAIGRKWRRSGDPFRC 111
 DB 1 VSEYSCRELHFTRYVTDGPGCRSAKPVTELVCSGCGPARLLPNAIGRKWRRSGDPFRC 60
 QY 112 IPDRYRQVRVQLCPGGEAPRARKVRLVASCCKRLTRFHNOSLKDFTGARPOKGRK 171
 DB 61 IPDRYRQVRVQLCPGGEAPRARKVRLVASCCKRLTRFHNOSLKDFTGARPOKGRK 120
 QY 172 PRPARSAKANQAELENAY 190
 DB 121 PRPARSAKANQAELENAY 139

RESULT 12
 ID AAB26105 standard; Protein; 367 AA.

XX AAB26105;
 AC AAB26105;

DT 15-JAN-2001 (first entry)

DE Human DAN/Cerberus-related protein 6 (hDCR6) #1.

XX Human; DNA/Cerberus-related protein 6; hDCR6; morphogenic protein;
 KW antagonist; BMP; cell growth; cell differentiation; bone formation;
 KM gene therapy.

XX Homo sapiens.

OS WO200055193-A2.

PN 21-SEP-2000.

PD 02-MAR-2000; 2000WO-US05537.

PF 12-MAR-1999; 99US-0124118.

PR (REGG-) REGENERON PHARM INC.

PA Economides AM;

PI WPI; 2000-638179/61.

DR N-PSDB; AAA94049, AAA94050.

XX Novel isolated, human DNA/Cerberus related protein 6 which include
 PT natural homologue, and polypeptides comprising DCR6 domain and nucleic
 PT acids encoding the proteins which are useful as probes and primers -
 XX Claim 7; Fig 2; 40pp; English.

XX The present sequence comprises the human DAN/Cerberus-related protein 6
 CC (hDCR6). Its coding sequence was isolated from a genomic DNA clone

CC following identification using computer-based 'virtual cloning'. hDCR6
 CC is closely related to the DAN and DCR5 proteins, both of which act as
 CC antagonists of morphogenic proteins such as BMP. It is possible that
 CC the hDCR6 gene and protein can be used as immunogens, modulators of cell
 CC function, growth and differentiation, to reduce undesirable bone
 CC formation, to identify DCR6 binding agents, in diagnosis, and in gene
 CC therapy.

XX Sequence 367 AA;

Query Match 73.2%; Score 139; DB 21; Length 367;
 Best Local Similarity 100.0%; Pred. No. 8e-137;
 Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 52 VSEYSCRELHFTRYVTDGPGCRSAKPVTELVCSGCGPARLLPNAIGRKWRRSGDPFRC 111
 DB 229 VSEYSCRELHFTRYVTDGPGCRSAKPVTELVCSGCGPARLLPNAIGRKWRRSGDPFRC 288
 QY 112 IPDRYRQVRVQLCPGGEAPRARKVRLVASCCKRLTRFHNOSLKDFTGARPOKGRK 171
 DB 289 IPDRYRQVRVQLCPGGEAPRARKVRLVASCCKRLTRFHNOSLKDFTGARPOKGRK 348
 QY 172 PRPARSAKANQAELENAY 190
 DB 349 PRPARSAKANQAELENAY 367

RESULT 13
 ID AAY96431 standard; Protein; 213 AA.

XX AAY96431;
 AC AAY96431;

DT 12-SEP-2000 (first entry)

DE Vervet TGF-beta binding protein (BEER).

XX osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
 KW BEER; gene therapy; antisense therapy; fracture; bone mineralization.

OS Cercopithecus pygerythrus.

PN WO200032773-A1.

PD 08-JUN-2000.

PF 24-NOV-1999; 99WO-US27990.

PR 27-NOV-1998; 98US-0110283.

PA (DARW-) DARWIN DISCOVERY LTD.

PI Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepker BW;

PT Van Ness J, Winkler DG;

DR WPI; 2000-412321/35.

DR N-PSDB; AAA29057.

PT Nucleic acids (1) encoding a transforming growth factor beta binding
 PT protein, useful for identifying agents for treating osteopenia,
 PT osteoporosis and fractures

PS Claim 4; Page 122-123; 162pp; English.

XX This shows a vervet transforming growth factor-beta (TGF-beta)
 CC binding protein designated VBER. The cDNA and protein may be used for
 CC prevention, treatment and diagnosis of diseases associated with
 CC inappropriate BEER expression. For example, they may be used to treat
 CC disorders associated with decreased TGF-beta BP expression. The cDNA or
 CC vectors may be administered to treat diseases by rectifying mutations or
 CC deletions in a patient's genome that affect the activity of BEER by
 CC expressing inactive proteins or to supplement the patients own production
 CC of BEER polypeptides. The nucleic acids may be used for recombinant

CC production of BEER, gene therapy, antisense therapy, as probes for
 CC diagnostic assays and for functional studies. BEER may be used to raise
 CC antibodies and for identification of BEER modulators. BEER antagonists
 CC may be used to increase bone mineral content for the treatment of
 CC disorders such as osteopenia, osteoporosis, fractures and other disorders
 CC associated with low mineral content.

XX Sequence 213 AA;

Query Match 67.4%; Score 128; DB 21; Length 213;
 Best Local Similarity 100.0%; Pred. No. 1.6e-125;
 Matches 128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 OQQAQKNDATETIIPPLGYPPEPELENNKTMNAENGRRPHHPFETKDVSEYSCREL 60
 DB 24 OQQAQKNDATETIIPPLGYPPEPELENNKTMNAENGRRPHHPFETKDVSEYSCREL 83
 QY 61 HFTRYVTDGPCRSAPKPYTELVCSCGCGPARLLPNAIGRGKWRPSPGDFRCI 120
 DB 84 HFTRYVTDGPCRSAPKPYTELVCSCGCGPARLLPNAIGRGKWRPSPGDFRCI 143
 QY 121 VOLLCPGG 128
 DB 144 VOLLCPGG 151

RESULT 14

AA96433
 ID AAY96434 standard; Protein; 176 AA.

AC AAY96434;

DT 12-SEP-2000 (first entry)

DE Bovine TGF-beta binding protein (BEER).

KM osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
 KW BEER; gene therapy; antisense therapy; fracture; bone mineralization.

OS Bos taurus.

PN WO200032773-A1.

PD 08-JUN-2000.

PF 24-NOV-1999; 99MO-US27990.

PR 27-NOV-1998; 98US-0110283.

PA (DARW-) DARWIN DISCOVERY LTD.

PI Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepfer BW;
 PI Van Ness J, Winkler DG;

DR WPI; 2000-412321/35.

DR N-PSDB; AAA29060.

PT Nucleic acids (I) encoding a transforming growth factor beta binding
 PT protein, useful for identifying agents for treating osteopenia,
 PT osteoporosis and fractures

PS Claim 7; Page 127; 162pp; English.

CC This shows a bovine transforming growth factor-beta (TGF-beta)
 CC binding protein designated bBEER. The cDNA and protein may be used for
 CC prevention, treatment and diagnosis of diseases associated with
 CC inappropriate BEER expression. For example, they may be used to treat
 CC disorders associated with decreased TGF-beta BP expression. The cDNA or
 CC vectors may be administered to treat diseases by rectifying mutations or
 CC deletions in a patient's genome that affect the activity of BEER by
 CC expressing inactive proteins or to supplement the activity of BEER by
 CC production of BEER, gene therapy, antisense therapy, as probes for

CC diagnostic assays and for functional studies. BEER may be used to raise
 CC antibodies and for identification of BEER modulators. BEER antagonists
 CC may be used to increase bone mineral content for the treatment of
 CC disorders such as osteopenia, osteoporosis, fractures and other disorders
 CC associated with low mineral content.

XX Sequence 176 AA;

Query Match 40.0%; Score 76; DB 21; Length 176;
 Best Local Similarity 100.0%; Pred. No. 3.4e-71;
 Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 53 SEYSCRELHFTRYVTDGPCRSAPKPYTELVCSCGCGPARLLPNAIGRGKWRPSPGDFRCI 112
 DB 45 SEYSCRELHFTRYVTDGPCRSAPKPYTELVCSCGCGPARLLPNAIGRGKWRPSPGDFRCI 104

QY 113 PDRYRAQRVOLLCPGG 128
 DB 105 PDRYRAQRVOLLCPGG 120

RESULT 15

AA96433
 ID AAY96433 standard; Protein; 213 AA.

AC AAY96433;

DT 12-SEP-2000 (first entry)

DE Rat TGF-beta binding protein (BEER).

KM osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
 KW BEER; gene therapy; antisense therapy; fracture; bone mineralization.

OS Rattus norvegicus.

PN WO200032773-A1.

PD 08-JUN-2000.

PF 24-NOV-1999; 99MO-US27990.

PR 27-NOV-1998; 98US-0110283.

PA (DARW-) DARWIN DISCOVERY LTD.

PI Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepfer BW;
 PI Van Ness J, Winkler DG;

DR WPI; 2000-412321/35.

DR N-PSDB; AAA29059.

PT Nucleic acids (I) encoding a transforming growth factor beta binding
 PT protein, useful for identifying agents for treating osteopenia,
 PT osteoporosis and fractures

PS Claim 6; Page 125-126; 162pp; English.

CC This shows a rat transforming growth factor-beta (TGF-beta) binding
 CC protein designated rBEER. The cDNA and protein may be used for
 CC prevention, treatment and diagnosis of diseases associated with
 CC inappropriate BEER expression. For example, they may be used to treat
 CC disorders associated with decreased TGF-beta BP expression. The cDNA or
 CC vectors may be administered to treat diseases by rectifying mutations or
 CC deletions in a patient's genome that affect the activity of BEER by
 CC expressing inactive proteins or to supplement the activity of BEER by
 CC production of BEER, gene therapy, antisense therapy, as probes for
 CC diagnostic assays and for functional studies. BEER may be used to raise
 CC antibodies and for identification of BEER modulators. BEER antagonists
 CC may be used to increase bone mineral content for the treatment of
 CC disorders such as osteopenia, osteoporosis, fractures and other disorders
 CC associated with low mineral content.

XX Sequence 213 AA;

Query Match 17.4%; Score 33; DB 21; Length 213;
 Best Local Similarity 100.0%; Pred. No. 4e-26;
 Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 66 VTGPGCRSAKPYTELVCSGCGGPARLLPNAIGR 98
 ||||||||||||||||||||||||||||||||
 Db 89 VTGPGCRSAKPYTELVCSGCGGPARLLPNAIGR 121

Search completed: March 28, 2003, 14:21:57
 Job time : 34.44 secs

GenCore version 5.1.4.P5_4576
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 28, 2003, 14:20:22 ; Search time 14.6933 Seconds

(without alignments)
1243.118 Million cell updates/sec

Title: US-09-867-274-2

Perfect score: 190

Sequence: 1 OGMQAFKNDATETIPELGEY.....KPRFRASAKANOALEINAY 190

Scoring table:

Gapop 60.0, Gapext 60.0

Searched: 283224 seqs, 96134422 residues

Word size: 0

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database:

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4.2	428	1	TVHUEK	transferring prote
2	4.2	429	2	JCA965	elk1 protein - mou
3	3.7	130	2	D8305	hypothetical prote
4	3.7	148	2	C83091	hypothetical prote
5	3.7	153	2	AE2284	hypothetical prote
6	3.7	183	2	B37410	H-2 class II histo
7	3.7	197	2	G81057	hypothetical prote
8	3.7	197	2	A81819	hypothetical prote
9	3.7	296	2	A12830	hypothetical prote
10	3.7	305	2	S77397	hypothetical prote
11	3.7	311	2	E98351	yabc protein homol
12	3.7	312	2	D75344	oligopeptide trans
13	3.7	321	2	B82892	conserved hypotet
14	3.7	328	2	S72647	hypothetical prote
15	3.7	341	2	AG2834	hypothetical prote
16	3.7	346	2	AB3323	3am-dependent meth
17	3.7	398	1	S24802	polyferredoxin 6x2
18	3.7	425	1	C89753	protein Flt7.3 [i
19	3.7	555	2	C97612	hypothetical prote
20	3.7	562	2	T05758	hypothetical prote
21	3.7	567	2	A87394	acyl-coa synthetas
22	3.7	602	2	S47880	NADH2 dehydrogenas
23	3.7	632	1	VGNVNSY	transcription anti
24	3.7	632	2	AC1547	transcription anti
25	3.7	667	2	UC7771	hypoxia inducible
26	3.7	681	2	B82812	outer membrane hem
27	3.7	781	1	TVEFDF	protein kinase Dya
28	3.7	917	1	S15885	hexokinase (EC 2.7
29	3.7	946	2	T16297	hypothetical prote

30	7	3.7	1008	2	B95979	probable enzyme wi
31	7	3.7	1488	2	AG2136	polyketide synthas
32	7	3.7	1940	2	A59287	myosin heavy chain
33	7	3.7	6420	2	T10283	polyketide synthas
34	6	3.2	49	2	S28215	neurotoxin Tx2 - s
35	6	3.2	65	2	B25025	malx protein - Kle
36	6	3.2	83	2	S41672	tightly associated
37	6	3.2	89	2	D75271	hypothetical prote
38	6	3.2	97	1	GECH	osteocalcin precur
39	6	3.2	102	2	G85584	unknown protein en
40	6	3.2	102	2	B90734	hypothetical prote
41	6	3.2	102	2	T51524	hypothetical prote
42	6	3.2	106	2	A82560	50S ribosomal prot
43	6	3.2	108	2	T30659	hypothetical prote
44	6	3.2	108	2	D90051	hypothetical prote
45	6	3.2	110	1	R5EC22	ribosomal protein

ALIGNMENTS

RESULT 1

TVHUEK

transferring protein elk-1 - human

C:Species: Homo sapiens (man)

C>Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 18-Jun-1999

C/Accession: A41354, S54721

R: Rao, V.N.; Huebner, K.; Isobe, M.; ar-Rushdi, A.; Croce, C.M.; Reddy, E.S.P.

Science 244, 66-70, 1989

A>Title: elk, tissue-specific ets-related genes on chromosomes X and 14 near translocated

A/Reference number: A41354; MUID:89203250; PMID:2539641

A/Accession: A41354

A/Molecule type: mRNA

A/Residues: 1-428 <RNO>

A/Cross-references: GB:M25269; NID:9538208; PIDN:AA52384.1; PID:G538209

R:Gille, H.; Kortelmann, M.; Thomae, O.; Moomaw, C.; Slaughter, C.; Cobb, M.H.; Shaw, P.I

EMBO J. 14, 951-962, 1995

A>Title: ERK phosphorylation potentiates Elk-1-mediated ternary complex formation and tr

A/Reference number: S54721; MUID:95196758; PMID:7889942

A/Accession: S54721

A/Status: preliminary

A/Molecule type: protein

A/Residues: 318-328, 'XX', 331, 336-364, 380-388, 'X', 390-392, 'X', 394-400, 'XX', 403-405, 'X', 407

C/Genetics:

A:Gene: GDB:ELK1

A:Cross-references: GDB:119867; OMIM:311040

A/Map position: Xp11.2-Xp11.2

C:Superfamily: elk-1,transferring protein; ets DNA-binding domain homology

C/Keywords: DNA binding; oncogene; transforming protein

F:7-86/Domain: ets DNA-binding domain homology <ETS>

Query Match

Best Local Similarity 4.2%; Score 8; DB 1; Length 428;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 166 POKRKPR 173

Db 310 POKRKPR 317

RESULT 2

JCA965

elk1 protein - mouse

C:Species: Mus musculus (house mouse)

C>Date: 31-Dec-1996 #sequence_revision 31-Dec-1996 #text_change 16-Jul-1999

C/Accession: JCA965; I48339; I48340; S54908

R:Grevin, D.; Ung, S.; Denhez, F.; Dehem, M.; Quatmann, B.; Begue, A.; Stehelin, D.; Me

Gene 174, 185-188, 1996

A>Title: Structure and organization of the mouse elk1 gene.

A/Reference number: JCA965; MUID:97017146; PMID:8863747

A/Accession: JCA965

A/Molecule type: mRNA

A/Residues: 1-429 <GRE>

A:Cross-references: EMBL:X87257; NID:g836634; PIDN:CAA60715.1; PID:g836635
 A:Experimental source: emdYto
 R:Giovane, A.; Pintzas, A.; Maira, S.M.; Sobieszczuk, P.; Wasyluk, B.
 Genes Dev. 8, 1502-1513, 1994
 A:Title: Net, a new ets transcription factor that is activated by Ras.
 A:Reference number: A53837; MUID:95047310; PMID:7958835
 A:Accession: 148339
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-429 <RES>
 A:Cross-references: EMBL:X87257; NID:g836634; PIDN:CAA60715.1; PID:g836635
 A:Accession: 148340
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 5-132, 'T', 134-224 <RES>
 A:Cross-references: EMBL:Z36939; NID:g535922; PIDN:CAA85391.1; PID:g535923
 C:Comment: This protein belongs to the subfamily of ternary complex factor (TCF) which h
 C:Genetics:
 A:Gene: elk1
 A:Introns: 70/3; 219/3; 363/3; 397/3
 C:Superfamily: elk-1 transforming protein; ets DNA-binding domain homology
 F:7-86/Domain: ets DNA-binding domain homology <ETS>
 Query Match 4.2%; Score 8; DB 2; Length 429;
 Best Local Similarity 100.0%; Pred. No. 3.6;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 166 POKGRKR 173
 DB 311 POKGRKR 318

RESULT 3
 D83305
 Hypothetical protein PA2722 [imported] - Pseudomonas aeruginosa (strain PA01)
 C:Species: Pseudomonas aeruginosa
 C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
 C:Accession: D83305
 A:Reference number: A82950; MUID:20437337; PMID:10984043
 A:Accession: D83305
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-130 <STO>
 A:Cross-references: GB:AE004700; GB:AE004091; NID:g9948792; PIDN:AG06110.1; GSPDB:GN001
 A:Experimental source: strain PA01
 C:Genetics:
 A:Gene: PA2722

Query Match 3.7%; Score 7; DB 2; Length 130;
 Best Local Similarity 100.0%; Pred. No. 16;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 135 KVRIVAS 141
 DB 12 KVRIVAS 18

RESULT 4
 C83091
 Hypothetical protein PA4441 [imported] - Pseudomonas aeruginosa (strain PA01)
 C:Species: Pseudomonas aeruginosa
 C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
 C:Accession: C83091
 R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; Lim,
 .; Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
 A:Reference number: A82950; MUID:20437337; PMID:10984043
 A:Accession: D83305
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-130 <STO>
 A:Cross-references: GB:AE004700; GB:AE004091; NID:g9948792; PIDN:AG06110.1; GSPDB:GN001
 A:Experimental source: strain PA01
 C:Genetics:
 A:Gene: PA2722

A:Reference number: A82950; MUID:20437337; PMID:10984043
 A:Accession: C83091
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-148 <STO>
 A:Cross-references: GB:AE004858; GB:AE004091; NID:g9950668; PIDN:AG07829.1; GSPDB:GN001;
 A:Experimental source: strain PA01
 C:Genetics:
 A:Gene: PA4441

Query Match 3.7%; Score 7; DB 2; Length 148;
 Best Local Similarity 100.0%; Pred. No. 17;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 89 ARLLPNA 95
 DB 26 ARLLPNA 32

RESULT 5
 AE2284
 Hypothetical protein alr3828 [imported] - Nostoc sp. (strain PCC 7120)
 C:Species: Nostoc sp.
 A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
 C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002
 C:Accession: AE2284
 R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,
 Nakazaki, N.; Shimo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
 DNA Res. 8, 205-213, 2001
 A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anal
 A:Reference number: AB1807; MUID:21595285; PMID:11758640
 A:Accession: AE2284
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-153 <KIR>
 A:Cross-references: GB:BA000019; PIDN:BA075527.1; PID:gl132962; GSPDB:GN00179
 A:Experimental source: strain PCC 7120
 C:Genetics:
 A:Gene: alr3828

Query Match 3.7%; Score 7; DB 2; Length 153;
 Best Local Similarity 100.0%; Pred. No. 18;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 9 DATEIIP 15
 DB 66 DATEIIP 72

RESULT 6
 B37410
 H-2 class II histocompatibility antigen A-2 beta chain - spiny mouse (Mus saxicola) (Frag
 C:Species: Mus saxicola
 C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 21-Jan-2000
 C:Accession: B37410
 R:Cam, P.; Jouvin-Marche, E.; LeGuern, C.; Marche, P.N.
 Eur. J. Immunol. 20, 1337-1343, 1990
 A:Title: Structure of class II genes in wild mouse Mus saxicola: functional and evolution
 A:Reference number: A37410; MUID:90316177; PMID:2369919
 A:Accession: B37410
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-183 <CAM>
 A:Cross-references: GB:M30158
 C:Superfamily: class II histocompatibility antigen; immunoglobulin homology
 F:105-170/Domain: immunoglobulin homology <IMM>
 Query Match 3.7%; Score 7; DB 2; Length 183;
 Best Local Similarity 100.0%; Pred. No. 20;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 12 EIIEIIG 18

Db 157 EITPELG 163

RESULT 7

G81057

hypothetical protein NMB1656 [imported] - *Neisseria meningitidis* (strain MCS8 serogroup C)Species: *Neisseria meningitidis*

C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001

C/Accession: G81057

R/Tettein, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.

Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;

Science 287, 1809-1815, 2000

A/Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappapoli, R.; V

A/Title: Complete genome sequence of *Neisseria meningitidis* serogroup B strain MCS8.

A/Reference number: A81000; PMID:20175755; PMID:10710307

A/Accession: G81057

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-197 <TER>

A/Cross-references: GB:AE002516; GB:AE002098; NID:G7226905; PIDN:AAF42005.1; PID:G722690

A/Experimental source: serogroup B, strain MCS8

C/Genetics:

A/Gene: NMB1656

Query Match

Best Local Similarity 100.0%; Pred. No. 22;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 180 KANQAEI 186

Db 122 KANQAEI 128

RESULT 8

A81819

hypothetical protein NMA1913 [imported] - *Neisseria meningitidis* (strain Z2491 serogroup C)Species: *Neisseria meningitidis*

C>Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001

C/Accession: A81819

R/Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel

Nature 404, 502-506, 2000

A/Title: Complete DNA sequence of a serogroup A strain of *Neisseria meningitidis* Z2491.

A/Reference number: A81775; PMID:20222556; PMID:10761919

A/Accession: A81819

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-197 <PAR>

A/Cross-references: GB:AL162757; GB:AL157959; NID:G7380371; PIDN:CAB85134.1; PID:G738054

A/Experimental source: serogroup A, strain Z2491

C/Genetics:

A/Gene: NMA1913

Query Match

Best Local Similarity 100.0%; Pred. No. 22;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 180 KANQAEI 186

Db 122 KANQAEI 128

RESULT 9

A12930

hypothetical protein Atu3047 [imported] - *Agrobacterium tumefaciens* (strain C58, Dupont)

C/Species: *Agrobacterium tumefaciens*

C/Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 01-Feb-2002

C/Accession: A12930

R/Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L

erage, G.; Gillit, W.; Grant, C.; Gentsch, D.; Kutyavin, T.; Levy, R.; Li, M.; McCell

Science 294, 2317-2323, 2001

A/Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, E

ster, E.W.

A/Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.

A/Reference number: AB2577; PMID:11743193

A/Accession: A12930

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-296 <KUN>

A/Cross-references: GB:AE006889; PIDN:AA43863.1; PID:G17741408; GSPDB:GN00187

A/Experimental source: strain C58 (Dupont)

C/Genetics:

A/Gene: Atu3047

A/Map position: linear chromosome

C/Superfamily: oligopeptide permease protein oppB

QY 91 LIPNIG 97

Db 206 LIPNIG 212

Query Match

Best Local Similarity 100.0%; Pred. No. 30;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 91 LIPNIG 97

Db 206 LIPNIG 212

RESULT 10

S77397

yabc protein homolog sl1144 - *Synechocystis* sp. (strain PCC 6803)

C/Species: *Synechocystis* sp.

A/Variety: PCC 6803

C/Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000

C/Accession: S77397

R/Kaneko, T.; Sato, S.; Korani, H.; Tanaka, A.; Aamizu, E.; Nakamura, Y.; Miyajima, N.

o. K.; Okumura, S.; Shimo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda,

DNA Res. 3, 109-136, 1996

A/Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis*

S.

A/Reference number: S74322; PMID:97061201; PMID:8905231

A/Accession: S77397

A/Status: nucleic acid sequence not shown; translation not shown.

A/Molecule type: DNA

A/Residues: 1-305 <KAN>

A/Cross-references: EMBL:D90906; GB:AB001339; NID:G1652492; PIDN:BA17500.1; PID:G1652575

A/Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

C/Superfamily: Escherichia coli yabc protein

Query Match

Best Local Similarity 100.0%; Pred. No. 31;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 174 PRARSAR 180

Db 292 PRARSAR 298

RESULT 11

E98351

oligopeptide transport system permease protein appC AGR_L_3516 [imported] - *Agrobacterium*

C/Species: *Agrobacterium tumefaciens*

C/Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 11-Jan-2002

C/Accession: E98351

R/Goodner, B.; Hinkle, G.; Gatung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,

Science 294, 2323-2328, 2001

A/Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent *Agrobacterium tum*

A/Reference number: A97359; PMID:11743194

A/Accession: E98351

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-311 <KUN>

A/Cross-references: GB:AE007870; PIDN:AAK90335.1; PID:G15160372; GSPDB:GN00170

C/Genetics:

A/Gene: AGR_L_3516

A/Map position: linear chromosome

C:Superfamily: oligopeptide permease protein oppb

Query Match 3.7%; Score 7; DB 2; Length 311;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 91 LLPNAIG 97
DB 221 LLPNAIG 227

RESULT 12

D75344 conserved hypothetical protein - Deinococcus radiodurans (strain R1)

C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
C:Accession: D75344
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; M.; Shen, M.; Vamathevan, J.D.; Lam, P.; McDonald, L.; Uterback, T.; Zaleski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999

A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.

A:Reference number: A75250; MUID:20036896; PMID:10567266

A:Accession: D75344

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-312 <NHI>

A:Cross-references: GB:AE002026; GB:AE000513; NID:G6459639; PIDN:AAF11419.1; PID:G645964

A:Experimental source: strain R1

C:Genetics:

A:Gene: DR1866

A:Map position: 1

C:Superfamily: Escherichia coli yabc protein

Query Match 3.7%; Score 7; DB 2; Length 312;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 174 PRARSAK 180
DB 293 PRARSAK 299

RESULT 13

B82892 conserved hypothetical UUA17 [imported] - Ureaplasma urealyticum

C:Species: Ureaplasma urealyticum
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000

C:Accession: B82892

R:Glaser, J.I.; Lefkowitz, E.J.; Glaser, J.S.; Heiner, C.R.; Chen, E.Y.; Caswell, G.H.

submitted to Genbank, February 2000

A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a mit

A:Reference number: A82870

A:Accession: B82892

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-321 <GLA>

A:Cross-references: GB:AE002139; GB:AF222894; NID:G6899405; PIDN:AAF10828.1; GSPDB:GN004

A:Experimental source: serovar 3; biovar 1

C:Genetics:

A:Gene: UUA17

A:Genetic code: SGC3

Query Match 3.7%; Score 7; DB 2; Length 321;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 135 KRVIVAS 141
DB 276 KRVIVAS 282

RESULT 14

572647 hypothetical protein 1b - antracnose fungus (Colleotrichum gloeosporioides) retrotransi
C:Species: Colleotrichum gloeosporioides
C:Date: 24-Oct-1998 #sequence_revision 24-Oct-1998 #text_change 17-Mar-1999
C:Accession: S72647
R:He, C.; Nourse, J.P.; Kelemu, S.; Irwin, J.A.G.; Manners, J.M.
Mol. Gen. Genet. 252, 320-331, 1996
A:Title: CgT1: a non-LTR retrotransposon with restricted distribution in the fungal phyc
A:Reference number: S72619; MUID:96439839; PMID:8842152
A:Accession: S72647
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-328 <HBH>
A:Experimental source: biotype B, isolate UQ62
C:Genetics:
A:Mobile element: retrotransposon CgT1

Query Match 3.7%; Score 7; DB 2; Length 328;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 87 GPARLLP 93
DB 123 GPARLLP 129

RESULT 15

AG2834 SAM-dependent methyltransferase [imported] - Agrobacterium tumefaciens (strain C58, Dupont

C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 01-Feb-2002

C:Accession: AG2834

R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Moo, L.

erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kuyavlin, T.; Levy, R.; Li, M.; McClellan

; Karp, P.; Romero, P.; Zhang, S.

Science 294, 2317-2323, 2001

A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, E

ster, E.W.

A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A:Reference number: AB2577; PMID:11743193

A:Accession: AG2834

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-341 <KUR>

A:Cross-references: GB:AE008688; PIDN:AAU43093.1; PID:G17740563; GSPDB:GN00186

A:Experimental source: strain C58 (Dupont)

C:Genetics:

A:Gene: maw

A:Map position: circular chromosome

C:Superfamily: Escherichia coli yabc protein

Query Match 3.7%; Score 7; DB 2; Length 341;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 174 PRARSAK 180
DB 303 PRARSAK 309

Search completed: March 28, 2003, 14:23:57
Job time: 15.6933 secs

DR MLM; 605740; -.
DB MTM; 368500; -.
DB

DR InterPro; IPR000359; Cys_knot.
 DR SMART; SMO0041; CT; 1.
 DR PROSITE; PS01185; CTCK_1; FALSE_NEG.
 DR PROSITE; PS01225; CTCK_2; FALSE_NEG.
 KW Signal; Glycoprotein.
 FT SIGNAL 1 23 POTENTIAL.
 FT CHAIN 24 213 SCLEROSTIN.
 FT DOMAIN 82 172 CTCK.
 FT CARBOHYD 53 53 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 175 175 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 213 AA; 24030 MW; 30DBD5CE73D5B2 CRC64;
 Query Match 100.0%; Score 190; DB 1; Length 213;
 Best Local Similarity 100.0%; Pred. No. 8.3e-195;
 Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 OGWOAFKNDATETIIPELGEYEPPEPELENNKTMNRAENGRRPHHPETKDVSEYSCREL 60
 DB 24 OGWOAFKNDATETIIPELGEYEPPEPELENNKTMNRAENGRRPHHPETKDVSEYSCREL 83
 QY 61 HFTRYVTDGPCRSAPVTELVCSGCCGPARLLPNAIGRGKWRPSPGDPFCIPDRYRAQR 120
 DB 84 HFTRYVTDGPCRSAPVTELVCSGCCGPARLLPNAIGRGKWRPSPGDPFCIPDRYRAQR 143
 QY 121 VQLCPGGAPARARVRLVASCCKRLTRFNQSELDFTGARPOKGRKPRPARSAK 180
 DB 144 VQLCPGGAPARARVRLVASCCKRLTRFNQSELDFTGARPOKGRKPRPARSAK 203
 QY 181 ANQAELENNAY 190
 DB 204 ANQAELENNAY 213
 RESULT 2
 SOST CERAE STANDARD; PRT; 213 AA.
 ID SOST CERAE
 AC 09BG78;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Sclerostin precursor.
 GN SOST.
 OS Cercopithecus aethiops (Green monkey) (Grivet).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecinae; Cercopithecus.
 NCBI_TaxID=9534;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21090529; PubMed=11179006;
 RA Brunkow M.E., Gardner J.C., Van Ness J., Paepker B.W., Kovacevich B.R.,
 RA Proll S., Skonier J.E., Zhao L., Sabo P.J., Fu Y.H., Alisch R.S.,
 RA Gillett L., Colbert T., Tacconi P., Galas D., Hamersma H.,
 RA Beighton P., Mulligan J.T.;
 RT "Bone dysplasia sclerosteosis results from loss of the SOST gene
 product, a novel cysteine knot-containing protein.";
 RL Am. J. Hum. Genet. 68:577-589(2001).
 CC -1- FUNCTION: Seems to play a role in bone homeostasis (By
 similarity).
 CC -1- SUBCELLULAR LOCATION: Secreted (Potential).
 CC -1- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTINE KNOT-LIKE (CTCK) DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC EMBL; AF326742; AAK13457.1; -
 DR InterPro; IPR000359; Cys_knot.
 DR SMART; SMO0041; CT; 1.

DR PROSITE; PS01185; CTCK_1; FALSE_NEG.
 DR PROSITE; PS01225; CTCK_2; FALSE_NEG.
 KW Signal; Glycoprotein.
 FT SIGNAL 1 23 POTENTIAL.
 FT CHAIN 24 213 SCLEROSTIN.
 FT DOMAIN 82 172 CTCK.
 FT CARBOHYD 53 53 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 175 175 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 213 AA; 23908 MW; 6DA7B5ED674728A CRC64;
 Query Match 67.4%; Score 128; DB 1; Length 213;
 Best Local Similarity 100.0%; Pred. No. 9.4e-129;
 Matches 128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 OGWOAFKNDATETIIPELGEYEPPEPELENNKTMNRAENGRRPHHPETKDVSEYSCREL 60
 DB 24 OGWOAFKNDATETIIPELGEYEPPEPELENNKTMNRAENGRRPHHPETKDVSEYSCREL 83
 QY 61 HFTRYVTDGPCRSAPVTELVCSGCCGPARLLPNAIGRGKWRPSPGDPFCIPDRYRAQR 120
 DB 84 HFTRYVTDGPCRSAPVTELVCSGCCGPARLLPNAIGRGKWRPSPGDPFCIPDRYRAQR 143
 QY 121 VQLCPGG 128
 DB 144 VQLCPGG 151
 RESULT 3
 SOST BOVIN STANDARD; PRT; 176 AA.
 ID SOST BOVIN
 AC 09BG79;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Sclerostin (Fragment).
 GN SOST.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21090529; PubMed=11179006;
 RA Brunkow M.E., Gardner J.C., Van Ness J., Paepker B.W., Kovacevich B.R.,
 RA Proll S., Skonier J.E., Zhao L., Sabo P.J., Fu Y.H., Alisch R.S.,
 RA Gillett L., Colbert T., Tacconi P., Galas D., Hamersma H.,
 RA Beighton P., Mulligan J.T.;
 RT "Bone dysplasia sclerosteosis results from loss of the SOST gene
 product, a novel cysteine knot-containing protein.";
 RL Am. J. Hum. Genet. 68:577-589(2001).
 CC -1- FUNCTION: Seems to play a role in bone homeostasis (By
 similarity).
 CC -1- SUBCELLULAR LOCATION: Secreted (Potential).
 CC -1- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTINE KNOT-LIKE (CTCK) DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC EMBL; AF326738; AAK13453.1; -
 DR InterPro; IPR000359; Cys_knot.
 DR SMART; SMO0041; CT; 1.
 DR PROSITE; PS01185; CTCK_1; FALSE_NEG.
 DR PROSITE; PS01225; CTCK_2; FALSE_NEG.
 KW Glycoprotein.
 FT DOMAIN 1 1 CTCK.
 FT NON TER 51 141 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 22 22

FT CARBOHYD 144 144 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT NON TER 176 176
 SQ SEQUENCE 176 AA; 19743 MW; 35FOCA61A425F4DB CRC64;

Query Match 40.0%; Score 76; DB 1; Length 176;
 Best Local Similarity 100.0%; Pred. No. 2e-73;
 Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 53 SEYSCREHAFTRVYTDGPCRSAPVTELVCSSGCGPARLLPNAIGKGMWRPSPDPFRCI 112
 DB 45 SEYSCREHAFTRVYTDGPCRSAPVTELVCSSGCGPARLLPNAIGKGMWRPSPDPFRCI 104

QY 113 PDYRAQRVQLCPGG 128
 DB 105 PDYRAQRVQLCPGG 120

RESULT 4

SOST_RAT STANDARD; PRT; 213 AA.
 ID SOST_RAT
 AC 099p67;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Sclerostin precursor.
 GN SOST.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;

RA Brunkow M.E., Gardner J.C., Van Ness J., Paepers B.W., Kovacevich B.R.,
 RA Prohl S., Skonier J.E., Zhao L., Sabo P.J., Fu Y.H., Alisch R.S.,
 RA Gillett L., Colbert T., Tacconi P., Galas D., Hamersma H.,
 RA Beighton P., Mulligan J.T.;
 RT "Bone dysplasia sclerosteosis results from loss of the SOST gene
 RT product, a novel cystine knot-containing protein.";
 RL Am. J. Hum. Genet. 68:577-589(2001).
 CC -1- FUNCTION: Seems to play a role in bone homeostasis (By
 CC similarity).
 CC -1- SUBCELLULAR LOCATION: Secreted (Potential).
 CC -1- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTINE KNOT-LIKE (CTCK) DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; AF326741; AAK13456.1; -;
 DR InterPro; IPR000359; Cys_knot.
 DR PROSITE; PS01185; CTCK_1; FALSE_NEG.
 DR PROSITE; PS01225; CTCK_2; FALSE_NEG.
 KW Signal; Glycoprotein.
 FT SIGNAL 1 28 POTENTIAL.
 FT CHAIN 29 213 SCLEROSTIN.
 FT DOMAIN 82 172 CTCK.
 FT CARBOHYD 53 53 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 175 175 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 213 AA; 23974 MW; 6C56C878C8D684B CRC64;

Query Match 17.4%; Score 33; DB 1; Length 213;
 Best Local Similarity 100.0%; Pred. No. 1.5e-27;
 Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 66 VTDGPCRSAPVTELVCSSGCGPARLLPNAIGR 98
 DB 89 VTDGPCRSAPVTELVCSSGCGPARLLPNAIGR 121

RESULT 5

SOST_MOUSE STANDARD; PRT; 211 AA.
 ID SOST_MOUSE
 AC 099p68; Q9D3L7;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Sclerostin precursor.
 GN SOST.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;

RA Brunkow M.E., Gardner J.C., Van Ness J., Paepers B.W., Kovacevich B.R.,
 RA Prohl S., Skonier J.E., Zhao L., Sabo P.J., Fu Y.H., Alisch R.S.,
 RA Gillett L., Colbert T., Tacconi P., Galas D., Hamersma H.,
 RA Beighton P., Mulligan J.T.;
 RT "bone dysplasia sclerosteosis results from loss of the SOST gene
 RT product, a novel cystine knot-containing protein.";
 RL Am. J. Hum. Genet. 68:577-589(2001).
 CC -1- FUNCTION: Seems to play a role in bone homeostasis (By
 CC similarity).
 CC -1- SUBCELLULAR LOCATION: Secreted (Potential).
 CC -1- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTINE KNOT-LIKE (CTCK) DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; AF326740; AAK13455.1; -;
 DR EMBL; AK017225; BAB30678.1; -;
 DR EMBL; AF326737; AAK13452.1; -;
 DR MGD; MGI:1921749; Sost.
 DR InterPro; IPR000359; Cys_knot.
 DR PROSITE; PS01185; CTCK_1; FALSE_NEG.
 DR PROSITE; PS01225; CTCK_2; FALSE_NEG.
 KW Signal; Glycoprotein.
 FT SIGNAL 1 23 POTENTIAL.
 FT CHAIN 24 211 SCLEROSTIN.

FT DOMAIN 80 170 CTCK.
 FT CARBOHYD 51 51 N-LINKED (GLCNAC. . .) (POTENTIAL)
 FT CARBOHYD 173 173 N-LINKED (GLCNAC. . .) (POTENTIAL)
 FT CONFLICT 72 72 G -> D (IN REF. 2).
 SQ SEQUENCE 211 AA; 23443 MW; AEB094E358B34961 CRC64;
 Query Match 16.8%; Score 32; DB 1; Length 211;
 Best Local Similarity 100.0%; Pred. No. 1.8e-26;
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 67 TDGCRSAKPVTELVCSGCCPARLLPNAIGR 98
 Db 88 TDGCRSAKPVTELVCSGCCPARLLPNAIGR 119
 RESULT 6
 ELK1_HUMAN STANDARD; PRT; 428 AA.
 AC P19419; O75606; O9UUM4; O95058;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE ETS-domain protein ELK-1.
 GN ELK1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89203250; PubMed=2539641;
 RA Rao V.N., Huebner K., Isobe M., Ar-Rushdi A., Croce C.M.,
 RA Reddy E.S.P.;
 RT "Elk, tissue-specific ets-related genes on chromosomes X and 14 near
 RT translocation breakpoints.";
 RL Science 244:66-70(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99013876; PubMed=9795224;
 RA Hardtman N., Mills F.C., Mitchell M.P., Meindl A., Max E.E.;
 RT "The human elk-1 gene family: the functional gene and two processed
 RT pseudogenes embedded in the Igh locus.";
 RL Gene 221:215-224(1998).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Graftiam D.;
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RA Ayee D.N.T., Kovar H.;
 RT "Novel family members HuER71, ELFR, and ELKV among ETS-related genes
 RT coexpressed with EMS-FLI1 in Ewing tumor cell lines.";
 RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP DOMAINS.
 RX MEDLINE=92334979; PubMed=1630903;
 RA Janknecht R., Nordheim A.;
 RT "Elk-1 protein domains required for direct and SRF-assisted
 RT DNA-binding.";
 RL Nucleic Acids Res. 20:3317-3324(1992).
 CC -1- FUNCTION: STIMULATES TRANSCRIPTION. BINDS TO PURINE-RICH DNA
 CC SEQUENCES. CAN FORM A TERNARY COMPLEX WITH THE SERUM RESPONSE
 CC FACTOR AND THE ETS AND SRF MOTIFS OF THE FOS SERUM RESPONSE
 CC ELEMENT.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 (SHOWN HERE) AND 2/ELKV; ARE
 CC PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: LUNG AND TESTIS.
 CC -1- PTM: PHOSPHORYLATED BY THE THREE GROUPS OF MAP KINASE (P38, JNK,
 CC AND ERK).
 CC -1- SIMILARITY: BELONGS TO THE ETS FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M25269; AAA52384.1; -;
 CC EMBL; AF080616; AAC82466.1; -;
 CC EMBL; AL009172; CAA15659.1; -;
 CC EMBL; AF000672; AAD00862.1; -;
 CC PIR; A41354; TVHUEK.
 CC HSSP; P28324; 1BC8.
 CC TRASNFAIC; T00250; -;
 CC Genew; HGNC:3321; ELK1.
 CC MIM; 311040; -;
 CC InterPro; IPR000418; Ets.
 CC InterPro; IPR002341; HSF_ETs.
 CC Pfam; PF00178; Ets; 1.
 CC PRINTS; PR00454; ETSDOMAIN.
 CC SMART; SM00413; ETS; 1.
 CC PROSITE; PS00345; ETS_DOMAIN_1; 1.
 CC PROSITE; PS00346; ETS_DOMAIN_2; 1.
 CC PROSITE; PS50061; ETS_DOMAIN_3; 1.
 CC KX Transcription regulation; Activator; Nuclear protein; DNA-binding;
 CC Phosphorylation; Alternative splicing.
 CC DNA_BIND 5 86 ETS-DOMAIN.
 CC VARSPIC 91 95 VAGCS -> SHCP (IN ISOFORM 2).
 CC VARSPLIC 96 428 MISSING (IN ISOFORM 2).
 CC CONFLICT 183 183 S -> N (IN REF. 1).
 SQ SEQUENCE 428 AA; 44888 MW; 68F71F8ADB9D38CA CRC64;
 Query Match 4.2%; Score 8; DB 1; Length 428;
 Best Local Similarity 100.0%; Pred. No. 1.2;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 166 PQGRKPR 173
 Db 310 PQGRKPR 317
 RESULT 7
 ELK1_MOUSE STANDARD; PRT; 429 AA.
 AC P41969;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE ETS-domain protein ELK-1.
 GN ELK1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Embryo;
 RX MEDLINE=97017146; PubMed=8663747;
 RA Grevin D., Ung S., Demuez F., Denem M., Quatamens B., Begue A.,
 RA Stenelin D., Martin P.;
 RT "Structure and organization of the mouse elk1 gene.";
 RL Gene 174:185-188(1996).
 RN [2]
 RP SEQUENCE OF 5-224 FROM N.A.
 RC TISSUE=Embryo;
 RX MEDLINE=95047310; PubMed=7958835;
 RA Giovane A., Pintzas A., Maira S.-M., Sobieszczuk P., Waeylyk B.;
 RT "Net, a new ets transcription factor that is activated by Ras.";
 RL Gene Dev. 8:1507-1513(1994).
 CC -1- FUNCTION: STIMULATES TRANSCRIPTION. BINDS TO PURINE-RICH DNA
 CC SEQUENCES. CAN FORM A TERNARY COMPLEX WITH THE SERUM RESPONSE
 CC FACTOR AND THE ETS AND SRF MOTIFS OF THE FOS SERUM RESPONSE
 CC ELEMENT.

CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN THE BRAIN, AND TO A
 CC LESSER EXTENT IN THE HEART, LIVER AND MUSCLE.
 CC -1- SIMILARITY: BELONGS TO THE ETS FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X67257; CAA60715.1; -
 CC EMBL: Z36939; CAA85391.1; -
 CC HSSP: P28324; IBC8.
 CC TRANSFAC: T05013; -
 CC MGD: MGI:101833; Elk1.
 CC InterPro: IPR000418; Ets.
 CC InterPro: IPR002341; HSF_ETs.
 CC Pfam: PF00178; Ets; 1.
 CC PRINTS: PR00454; ETSDOMAIN.
 CC SMART: SM00413; ETS; 1.
 CC PROSITE: PS00345; ETS_DOMAIN_1; 1.
 CC PROSITE: PS00346; ETS_DOMAIN_2; 1.
 CC PROSITE: PS50061; ETS_DOMAIN_3; 1.
 CC Transcription regulation; Activator; Nuclear protein; DNA-binding;
 KW Phosphorylation.
 KW DNA BIND
 FT CONFLICT 133 133 P -> T (IN REF. 2).
 SQ SEQUENCE 429 AA; 45243 MW; B61B5B97731054F CRC64;
 Query Match 4.2%; Score 8; DB 1; Length 429;
 Best Local Similarity 100.0%; Pred. No. 1.2;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 166 POKGRKRP 173
 Db 311 POKGRKRP 318
 RESULT 8
 RL22_LEPIN
 ID RL22_LEPIN STANDARD; PRT; 110 AA.
 AC Q9XD31;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE 50S ribosomal protein L22.
 GN RPLV.
 OS Leptospira interrogans.
 CC Bacteria; Spirochaetales; Leptospiraceae; Leptospira.
 CC NCBI_TaxId=173;
 CC [1]
 CC SEQUENCE FROM N.A.
 CC STRAIN=Setovar Lai;
 CC MEDLINE=20088835; PubMed=10620683;
 CC Zuercher R.L., Hartskeerl R.A., van de Kemp H., Bal A.3.;
 CC "Characterization of the Leptospira interrogans S10-spcc-alpha
 CC operon."
 CC FEMS Microbiol. Lett. 182:303-308(2000).
 CC -1- FUNCTION: THIS PROTEIN BINDS SPECIFICALLY TO 23S rRNA; ITS BINDING
 CC IS STIMULATED BY OTHER RIBOSOMAL PROTEINS, E.G., 24, L17, AND L20.
 CC IT IS IMPORTANT DURING THE EARLY STAGES OF 50S RECONSTITUTION
 CC (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE L22P FAMILY OF RIBOSOMAL PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch.
 CC -----
 CC EMBL: AF15283; AAD40588.1; -
 CC HSSP: P48286; IEXE.
 CC InterPro: IPR001063; Ribosomal_L22.
 CC Pfam: PF00237; Ribosomal_L22; 1.
 CC ProDom: PD001032; Ribosomal_L22; 1.
 CC TIGRFAMs: TIGR01044; rplV_bact; 1.
 CC PROSITE: PS00464; RIBOSOMAL_L22; 1.
 KW Ribosomal protein; rRNA-binding
 SQ SEQUENCE 110 AA; 12566 MW; B9181B85E9756CAA CRC64;
 Query Match 3.7%; Score 7; DB 1; Length 110;
 Best Local Similarity 100.0%; Pred. No. 4.3;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 134 RYRRLVA 140
 Db 15 RYRRLVA 21
 RESULT 9
 MRW_SYNY3
 ID MRW_SYNY3 STANDARD; PRT; 305 AA.
 AC P73460;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE S-adenosyl-methyltransferase mrw (EC 2.1.1.-).
 GN MRW OR SL1144.
 OS Synechocystis sp. (strain PCC 6803).
 CC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
 CC NCBI_TaxId=1148;
 CC [1]
 CC SEQUENCE FROM N.A.
 CC MEDLINE=97061201; PubMed=8905231;
 CC Kaneko T., Sato S., Kocani H., Tanaka A., Asamizu E., Nakamura Y.,
 CC Miyajima N., Hirosewa M., Sugitara M., Sasamoto S., Kimura T.,
 CC Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naito K., Okumura S.,
 CC Shimpo S., Takeuchi C., Wada T., Watanabe A., Yanada M., Yasuda M.,
 CC Tabata S.;
 CC "Sequence analysis of the genome of the unicellular cyanobacterium
 CC Synechocystis sp. strain PCC6803. II. Sequence determination of the
 CC RT entire genome and assignment of potential protein-coding regions."
 CC RL DNA Res. 3:109-136(1996).
 CC -1- FUNCTION: Exhibits a S-adenosyl-dependent methyltransferase
 CC activity (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE MRW FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: D90906; BAA17500.1; -
 CC InterPro: IPR002903; Bac_Metnfrse.
 CC Pfam: PF01795; Methyltransf_5; 1.
 CC ProDom: PD004685; Bac_Metnfrse; 1.
 CC TIGRFAMs: TIGR00006; UPF0117; 1.
 KW Transferase; Methyltransferase; Complete proteome.
 SQ SEQUENCE 305 AA; 34386 MW; 758F9F567247129F CRC64;
 Query Match 3.7%; Score 7; DB 1; Length 305;
 Best Local Similarity 100.0%; Pred. No. 10;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 174 PRASAK 180
 Db 292 PRASAK 298

```

RESULT 10
MRAM DEIRA          STANDARD;      PRT;      312 AA.
ID MRAM DEIRA
AC O9RT99;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE S-adenosyl-methyltransferase mraw (EC 2.1.1.-).
GN MRAW OR DR1866.
OS Deinococcus radiodurans.
OC Bacteria; Thermus/Deinococcus group; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Bisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zalewski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome sequence of the radioresistant bacterium Deinococcus
RT radiodurans R1.";
RL Science 286:1571-1577(1999).
CC -!- FUNCTION: Exhibits a S-adenosyl-dependent methyltransferase
CC activity (By similarity).
CC -!- SIMILARITY: BELONGS TO THE MRAW FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE002026; AAF1419.1; -.
DR TIGR; DR1866; -.
DR InterPro; IPR002903; Bac_MetTrfse.
DR Pfam; PF01795; Methyltransf_5; 1.
DR ProDom; PD004648; Bac_MetTrfse; 1.
DR TIGRFAMs; TIGR00006; UPP0117; 1.
KW Transferase; Methyltransferase; Complete proteome.
SQ SEQUENCE 312 AA; 33443 MW; B51835F2273A9463 CRC64;

Query Match          3.7%; Score 7; DB 1; Length 312;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 174 PRARSAX 180
Db 293 PRARSAX 299

RESULT 11
MRAM AGRT5          STANDARD;      PRT;      341 AA.
ID MRAM AGRT5
AC P58745;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE S-adenosyl-methyltransferase mraw (EC 2.1.1.-).
GN MRAW OR ATU2102 OR AGR C 3815.
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RX MEDLINE=21608550; PubMed=11743193;
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Boyce D. Sr.,
RA Chapman P., Clendenning J., Decherage G., Gillet W., Grant C.,
RA Kutayavin T., Levy R., Li M.-O., McClelland E., Palmeri A., Gordon D.,
RA Raymond C., Rouse G., Saenphithmachak C., Wu Z., Romero P., Perry M.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley P., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Neeter E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
RT C58.";
RL Science 294:2317-2323 (2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608551; PubMed=11743194;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M., Mullin L.,
RA Quirillo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
RA Houniel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
RA Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,
RA Flanagan C., Crowell C., Garson J., Lomo C., Sear C., Strub G.,
RA Cleto C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent
RT Agrobacterium tumefaciens C58.";
RL Science 294:2323-2328(2001).
CC -!- FUNCTION: Exhibits a S-adenosyl-dependent methyltransferase
CC activity (By similarity).
CC -!- SIMILARITY: BELONGS TO THE MRAW FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE009160; AAL43093.1; -.
DR EMBL; AE008126; AAK87852.1; ALT INIT.
DR InterPro; IPR002903; Bac_MetTrfse.
DR Pfam; PF01795; Methyltransf_5; 1.
KW Transferase; Methyltransferase; Complete proteome.
SQ SEQUENCE 341 AA; 36369 MW; 5EED5F0731B8C208 CRC64;

Query Match          3.7%; Score 7; DB 1; Length 341;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 174 PRARSAX 180
Db 303 PRARSAX 309

RESULT 12
MRAM RHIME          STANDARD;      PRT;      341 AA.
ID MRAM RHIME
AC O92N14;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE S-adenosyl-methyltransferase mraw (EC 2.1.1.-).
GN MRAW OR R02184 OR SMC01858.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=1021;
RX MEDLINE=21396507; PubMed=11481430;
RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Bactut J.,

```

RA Boileard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
 RA Godtke T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masny D.,
 RA Pohl T., Pothier D., Puehler A., Purnelle B., Ramsperger U.,
 RA Renard C., Thebaud P., Vandenbol M., Weidner S., Galibert F.,
 RT "Analysis of the chromosome sequence of the legume symbiont
 RT *Sinorhizobium meliloti* strain 1021.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
 CC -I- FUNCTION: Exhibits a S-adenosyl-dependent methyltransferase
 CC activity (By similarity).
 CC -I- SIMILARITY: BELONGS TO THE MRAM FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AL591789; CAC46763.1; -;
 DR InterPro; IPR002903; Bac_Metrifrase.
 DR Pfam; PF01795; Methyltransf_5; 1.
 DR ProDom; PD004685; Bac_Metrifrase; 1.
 DR TIGRPFAMs; TIGR00006; UPF0117; 1.
 KW Transferase; Methyltransferase; Complete proteome.
 SQ SEQUENCE 341 AA; 36021 MW; D84227F3D9B87938 CRC64;
 QY Query Match 3.7%; Score 7; DB 1; Length 341;
 DB Best Local Similarity 100.0%; Pred. No. 11;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 174 PRARSAK 180
 DB 303 PRARSAK 309
 ID MRAM BRUME STANDARD; PRT; 346 AA.
 AC Q8Y174;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE S-adenosyl-methyltransferase mraw (EC 2.1.1.-).
 GN MRAW OR BMEI0571.
 OS Brucella melitensis.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 CC Brucellaceae; Brucella.
 CC NCBI_TaxId=29459;
 CC (1)
 RN SEQUENCE FROM N.A.
 RC STRAIN=16M / ATCC 23456 / Biotype 1;
 RX MEDLINE=20020109; PubMed=11756688;
 RA Delbecch V.G., Kapral V., Redkar R.J., Patra G., Mujer C., Los T.,
 RA Ivanova N., Andersen T., Bhattacharya A., Lykidis A., Resnik G.,
 RA Jablonka L., Larsen N., D'Souza M., Bernal A., Mazur M., Goldsman E.,
 RA Selkov E., Elser P.H., Hagius S., O'Callaghan D., Lelsson J.-J.,
 RA Haselkorn R., Kyriades N., Overbeek R.;
 RT "The genome sequence of the facultative intracellular pathogen
 RT *Brucella melitensis*.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
 CC -I- FUNCTION: Exhibits a S-adenosyl-dependent methyltransferase
 CC activity (By similarity).
 CC -I- SIMILARITY: BELONGS TO THE MRAM FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; AE009499; AL51752.1; -;
 DR InterPro; IPR002903; Bac_Metrifrase.
 DR Pfam; PF01795; Methyltransf_5; 1.
 DR ProDom; PD004685; Bac_Metrifrase; 1.
 DR TIGRPFAMs; TIGR00006; UPF0117; 1.
 KW Transferase; Methyltransferase; Complete proteome.
 SQ SEQUENCE 346 AA; 37501 MW; 23F15D1F05D7B86 CRC64;
 QY Query Match 3.7%; Score 7; DB 1; Length 346;
 DB Best Local Similarity 100.0%; Pred. No. 12;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 174 PRARSAK 180
 DB 307 PRARSAK 313
 ID NUSM DIDMA STANDARD; PRT; 602 AA.
 AC P41309; P92474;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE NADH-ubiquinone oxidoreductase chain 5 (EC 1.6.5.3).
 GN MTRD5 OR ND5.
 OS Didelphis marsupialis virginiana (North American opossum).
 CC Mitochondrion.
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Metatheria; Didelphimorphia; Didelphidae; Didelphis.
 CC NCBI_TaxId=9267;
 CC (1)
 RN SEQUENCE FROM N.A.
 RP TISSUE=Liver.
 RC MEDLINE=94333786; PubMed=8056314;
 RA Janke A., Feldmaier-Fuchs G., Thomas K., von Haeseler A., Pabo S.;
 RT "The marsupial mitochondrial genome and the evolution of placental
 RT mammals.";
 RL Genetics 137:243-256(1994).
 CC -I- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Z29573; CAA82687.1; -;
 DR PIR; S42820; S42820.
 DR InterPro; IPR003916; NADH_oxred5.
 DR InterPro; IPR001516; Oxidored_q1_N.
 DR Pfam; PF00361; Oxidored_q1_N; 1.
 DR Pfam; PF00662; Oxidored_q1_N; 1.
 DR PRINTS; PR01434; NADHGHGASES.
 KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
 SQ SEQUENCE 602 AA; 68001 MW; 1020P172592669RE CRC64;
 QY Query Match 3.7%; Score 7; DB 1; Length 602;
 DB Best Local Similarity 100.0%; Pred. No. 19;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 27 LENKTM 33
 DB 267 LENKTM 273
 ID VGLG SYNK STANDARD; PRT; 632 AA.
 AC P21277;
 RESULT 15
 VGLG SYNK
 ID VGLG SYNK STANDARD; PRT; 632 AA.
 AC P21277;

```

DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-JUL-1993 (Rel. 26, Last annotation update)
DE Spike glycoprotein precursor.
GN g.
OS Sonchus yellow net virus (SYNV).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Rhabdoviridae; Nucleorhabdovirus.
OX NCBI_TaxID=11307;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC PV-263;
RX MEDLINE=92024089; PubMed=1926779;
RA Goldberg K.B., Modrell B., Hillman B.I., Heaton L.A., Choi T.J.,
RA Jackson A.O.;
RT "Structure of the glycoprotein gene of sonchus yellow net virus, a
RT plant rhabdovirus.";
RL Virology 185:32-38(1991).
CC -1- FUNCTION: THIS PROTEIN FORMS SPIKES ON THE SURFACE OF THE VIRION.
CC IT IS RESPONSIBLE BOTH FOR THE BINDING OF THE VIRUS TO SUSCEPTIBLE
CC HOST CELLS AND FOR INDUCING THE UPTAKE OF THE VIRUS BY THE CELL.
CC THE INTERACTION BETWEEN THE INTERNAL COMPONENTS OF THE VIRION
CC AND THE PORTION OF THE GLYCOPROTEIN EXPOSED ON THE CYTOPLASMIC
CC FACE OF THE PLASMA MEMBRANE PROBABLY DIRECTS ENVELOPMENT AND
CC VIRUS BUDDING.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L32603; AAA50384.1; -.
DR EMBL; M73626; AAA47898.1; -.
DR PIR; A40776; VGVNSY.
KW Transmembrane; Envelope protein; Glycoprotein; Signal.
FT STGNAL 1 17 POTENTIAL.
FT CHAIN 18 632 SPIKE GLYCOPROTEIN.
FT TRANSMEM 560 578 POTENTIAL.
FT CARBOHYD 93 93 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 385 385 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 501 501 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 512 512 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 541 541 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 632 AA; 71093 MW; 05541E78E8A07927 CRC64;

```

```

Query Match 3.7%; Score 7; DB 1; Length 632;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 135 KVRIVAS 141
DB 579 KVRIVAS 585

```

Search completed: March 28, 2003, 14:22:21
Job time : 9.10667 secs

GenCore version 5.1.4.P5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 28, 2003, 14:19:52 ; Search time 26.8533 Seconds
(without alignments)
1457.881 Million cell updates/sec

Title: US-09-867-274-2

Perfect score: 190
Sequence: 1 QGMOAFKNDATETIPELGEY.....KPRPRASAKAOALEENAY 190

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 671580 seqs, 206047115 residues

Word size: 0

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database:

SPTREMBL_21:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriophage:*
- 17: sp_archaeal:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	9	4.7	206	4 Q96HJ7	Q96HJ7 homo sapien
2	9	4.7	206	11 Q96HJ7	Q96HJ7 mus musculu
3	9	4.7	206	11 Q96HJ7	Q96HJ7 oryza sativ
4	8	4.2	276	13 Q9YGD6	Q9YGD6 oncorhynch
5	8	4.2	276	13 Q9PT36	Q9PT36 oncorhynch
6	8	4.2	276	13 Q9PT37	Q9PT37 oncorhynch
7	8	4.2	276	13 Q9PT38	Q9PT38 oncorhynch
8	8	4.2	428	4 Q965X8	Q965X8 homo sapien
9	8	4.2	89	10 Q9FH55	Q9FH55 arabidopsis
10	7	3.7	109	5 Q9GTU4	Q9GTU4 ancylostoma
11	7	3.7	130	16 Q9I0C0	Q9I0C0 pseudomonas
12	7	3.7	148	16 Q9HVK4	Q9HVK4 pseudomonas
13	7	3.7	153	16 Q9YQJ5	Q9YQJ5 anabaena sp
14	7	3.7	183	7 Q31243	Q31243 mus saxicol
15	7	3.7	186	10 Q9SDR7	Q9SDR7 forsythia i
16	7	3.7	197	16 Q9YB89	Q9YB89 neisseria m

17	7	3.7	197	16 Q9JTA0	Q9JTA0 neisseria m
18	7	3.7	214	10 Q8RUS7	Q8RUS7 oryza sativ
19	7	3.7	251	16 Q9LINO	Q9LINO streptomyces
20	7	3.7	282	10 Q9XFL9	Q9XFL9 arabidopsis
21	7	3.7	311	16 Q8UBG9	Q8UBG9 agrobacteri
22	7	3.7	321	16 Q9F072	Q9F072 ureaplasma
23	7	3.7	367	11 Q9IYE3	Q9IYE3 mus musculu
24	7	3.7	398	11 Q00388	Q00388 methanococc
25	7	3.7	425	5 Q19348	Q19348 caenorhabdi
26	7	3.7	552	10 Q81837	Q81837 arabidopsis
27	7	3.7	567	16 Q9A929	Q9A929 caulobacter
28	7	3.7	604	5 Q44003	Q44003 toxoplasma
29	7	3.7	632	4 Q8WXA1	Q8WXA1 homo sapien
30	7	3.7	632	4 Q96K34	Q96K34 homo sapien
31	7	3.7	632	16 Q92D98	Q92D98 listeria in
32	7	3.7	638	4 Q9H7Y7	Q9H7Y7 homo sapien
33	7	3.7	648	4 Q9HAI2	Q9HAI2 homo sapien
34	7	3.7	667	4 Q9Y2N7	Q9Y2N7 homo sapien
35	7	3.7	681	16 Q9PGB8	Q9PGB8 xylella fas
36	7	3.7	739	5 Q9BSC4	Q9BSC4 homo sapien
37	7	3.7	739	5 Q9W4Z3	Q9W4Z3 drosophila
38	7	3.7	782	5 Q9NEH9	Q9NEH9 drosophila
39	7	3.7	802	5 Q96398	Q96398 schistosoma
40	7	3.7	815	4 Q43273	Q43273 homo sapien
41	7	3.7	917	11 Q54892	Q54892 rattus norv
42	7	3.7	962	17 Q8TQC1	Q8TQC1 methanosarc
43	7	3.7	974	5 Q20143	Q20143 caenorhabdi
44	7	3.7	980	4 Q9NS55	Q9NS55 homo sapien
45	7	3.7	1008	16 Q9ZUM7	Q9ZUM7 rhizobium m

ALIGNMENTS

RESULT 1

Q96HJ7 PRELIMINARY; PRT; 206 AA.

AC Q96HJ7;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Similar to RIKEN cDNA 061006G05 gene (CDA019).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BONE MARROW;
RA Strauberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Liu F., Xu X.R., Qian B.Z., Xiao H., Chen Z., Han Z.;
RT "A novel gene expressed in human pheochromocytoma";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC008484; AA08484.1; -
DR EMBL; AF361494; AAL57219.1; -
DR InterPro; IPR000359; Cys_knot.
DR PROSITE; PS01225; CTCK_2; 1.
SQ SEQUENCE 206 AA; 23306 MW; 9FBJCC41E4B53834 CRC64;

Query Match 4.7%; Score 9; DB 4; Length 206;
Best Local Similarity 100.0%; Pred. No. 0.41;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 AFKNDATETI 13
Db 23 AFKNDATETI 31

RESULT 2
Q9CON4

```

ID 09CON4 PRELIMINARY; PRT; 206 AA.
AC 09CON4;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE 0610006G05RIK protein (RIKEN CDNA 0610006G05 gene).
GN 0610006G05RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CS7BL/6J; TISSUE=PANCREAS AND KIDNEY;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinaawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischnann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Knehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schirml L.M., Staudl F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustinich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli U., Mombauts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Wetz C., Whitaker C., Wilmig L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.,
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=KIDNEY;
RA Strausberg R.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK007967; BAB25378.1; -
DR EMBL; AK002240; BAB21957.1; -
DR EMBL; AK002396; BAB22068.1; -
DR EMBL; AK007693; BAB25333.1; -
DR EMBL; BC021458; AAR21458.1; -
DR WGD; MG11913292; 0610006G05RIK.
DR InterPro; IPR000359; Cys_knot.
DR PROSITE; PS01225; CTCK_2; 1.
SQ SEQUENCE 206 AA; 23174 MW; 70D24819EB06CBC9 CRC64;

Query Match 4.7%; Score 9; DB 11; Length 206;
Best Local Similarity 100.0%; Pred. No. 0.41;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 5 AFKDATEI 13
DB 23 AFKDATEI 31

RESULT 3
Q9FY03 PRELIMINARY; PRT; 198 AA.
AC Q9FY03;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Hypothetical protein (P0011G08.32 protein).
GN P0011G08.32.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Eubartoideae; Oryzaceae; Oryza.
OC NCBI_TaxID=4530;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nippobare (GA3) genomic DNA, chromosome 1, PAC
RT clone:P0433F09.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. MATSUMOTO; Yamamoto K.;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nippobare (GA3) genomic DNA, chromosome 1, PAC
RT clone:P0011G08.";
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP002553; BAB08182.1; -
DR EMBL; AP003225; BAB64660.1; -
DR InterPro; IPR000886; ER target.
DR PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
KW Hypothetical protein.
SQ SEQUENCE 198 AA; 21843 MW; 11DD90FE4B72F4E CRC64;

Query Match 4.2%; Score 8; DB 10; Length 198;
Best Local Similarity 100.0%; Pred. No. 4.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 126 PGGEAPRA 133
DB 83 PGGEAPRA 90

RESULT 4
Q9YGD6 PRELIMINARY; PRT; 276 AA.
AC Q9YGD6;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Carboxyl reductase/20beta-hydroxysteroid dehydrogenase B.
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Proacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=8022;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99185307; PubMed=10082666;
RA Guan G., Tanaka M., Todo T., Young G., Yoshikuni M., Nagahama Y.;
RT "Cloning and expression of two carboxyl reductase-like 20beta-
RT hydroxysteroid dehydrogenase cDNAs in ovarian follicles of rainbow
RT trout.";
RL Biochem. Biophys. Res. Commun. 255:123-128(1999).
CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
CC (SDR) FAMILY.
DR EMBL; AF100930; AAD20217.1; -
DR InterPro; IPR002198; ADH short.
DR Pfam; PF00106; adh short; 1.
DR PRINTS; PR00080; SDRFAMILY.
KW Oxidoreductase.
SQ SEQUENCE 276 AA; 30111 MW; 135604A38D2940CD CRC64;

Query Match 4.2%; Score 8; DB 13; Length 276;
Best Local Similarity 100.0%; Pred. No. 5.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 5 AFKDATEI 12
DB 93 AFKDATEI 100

RESULT 5
Q9PT36 PRELIMINARY; PRT; 276 AA.
AC Q9PT36;

```

DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Carboxyl reductase/20beta-hydroxysteroid dehydrogenase A.
 OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 CC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
 NCBI_TaxID=8022;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99185307; PubMed=10082666;
 RA Guan G., Tanaka M., Todo T., Young G., Yoshikuni M., Nagahama Y.;
 RT "Cloning and expression of two carboxyl reductase-like 20beta-
 RT hydroxysteroid dehydrogenase cDNAs in ovarian follicles of rainbow
 RT trout.";
 RL Biochem. Biophys. Res. Commun. 255:123-128(1999).
 CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
 CC (SDR) FAMILY.
 DR EMBL; AF100931; AAD20218.1; -.
 DR HSSP; P50162; IAE1.
 DR InterPro; IPR002198; ADH_short.
 DR Pfam; PF00106; adh_short; 1.
 DR PRINTS; PR00080; SDRFAMILY.
 KW Oxidoreductase.
 SQ SEQUENCE 276 AA; 30139 MW; 135F1711D30086D CRC64;

Query Match 4.2%; Score 8; DB 13; Length 276;
 Best Local Similarity 100.0%; Pred. No. 5.8;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 AFKNDATE 12
 DB 93 AFKNDATE 100

RESULT 6
 Q9PT37 PRELIMINARY; PRT; 276 AA.

DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Carboxyl reductase/20beta-hydroxysteroid dehydrogenase B.
 OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 CC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
 NCBI_TaxID=8022;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99185307; PubMed=10082666;
 RA Guan G., Tanaka M., Todo T., Young G., Yoshikuni M., Nagahama Y.;
 RT "Cloning and expression of two carboxyl reductase-like 20beta-
 RT hydroxysteroid dehydrogenase cDNAs in ovarian follicles of rainbow
 RT trout.";
 RL Biochem. Biophys. Res. Commun. 255:123-128(1999).
 CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
 CC (SDR) FAMILY.
 DR EMBL; AF100931; AAD20991.1; -.
 DR InterPro; IPR002198; ADH_short.
 DR Pfam; PF00106; adh_short; 1.
 DR PRINTS; PR00080; SDRFAMILY.
 KW Oxidoreductase.
 SQ SEQUENCE 276 AA; 30127 MW; 1350B4211D34225D CRC64;

Query Match 4.2%; Score 8; DB 13; Length 276;
 Best Local Similarity 100.0%; Pred. No. 5.8;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 AFKNDATE 12
 DB 93 AFKNDATE 100

RESULT 7
 Q9PT38 PRELIMINARY; PRT; 276 AA.
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Carboxyl reductase/20beta-hydroxysteroid dehydrogenase A.
 OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 CC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
 NCBI_TaxID=8022;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99185307; PubMed=10082666;
 RA Guan G., Tanaka M., Todo T., Young G., Yoshikuni M., Nagahama Y.;
 RT "Cloning and expression of two carboxyl reductase-like 20beta-
 RT hydroxysteroid dehydrogenase cDNAs in ovarian follicles of rainbow
 RT trout.";
 RL Biochem. Biophys. Res. Commun. 255:123-128(1999).
 CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
 CC (SDR) FAMILY.
 DR EMBL; AF100931; AAD20992.1; -.
 DR HSSP; P50162; IAE1.
 DR InterPro; IPR002198; ADH_short.
 DR Pfam; PF00106; adh_short; 1.
 DR PRINTS; PR00080; SDRFAMILY.
 KW Oxidoreductase.
 SQ SEQUENCE 276 AA; 30239 MW; 134B52054C751C28 CRC64;

Query Match 4.2%; Score 8; DB 13; Length 276;
 Best Local Similarity 100.0%; Pred. No. 5.8;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 AFKNDATE 12
 DB 93 AFKNDATE 100

RESULT 8
 Q969X8 PRELIMINARY; PRT; 428 AA.

DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE ELK1 protein.
 GN ELK1.
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=HIPPOCAMPUS.
 RX MEDLINE=99246057; PubMed=10231026;
 RA Yamauchi T., Toko M., Suga M., Hatakeyama T., Isebe M.;
 RT "Structural organization of the human ELK1 gene and its processed
 RT pseudogene ELK2 genes.";
 RL DNA Res. 6:21-27(1999).
 DR EMBL; AB016194; BAA36617.1; -.
 DR EMBL; AB016193; BAA36616.1; -.
 DR InterPro; IPR000418; Ets.
 DR InterPro; IPR002341; HSF_ETs.
 DR Pfam; PF00178; Ets; 1.
 DR PROSITE; PS00345; ETS_DOMAIN_1; UNKNOWN_1.
 DR PROSITE; PS00346; ETS_DOMAIN_2; UNKNOWN_1.
 DR PROSITE; PS50061; ETS_DOMAIN_3; 1.
 SQ SEQUENCE 428 AA; 44915 MW; 58FBFBAD9B985A83 CRC64;

Query Match 4.2%; Score 8; DB 4; Length 428;
 Best Local Similarity 100.0%; Pred. No. 8.4;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 166 POKGRKPR 173
 |||||
 DB 310 POKGRKPR 317

RESULT 9

O9FH55 PRELIMINARY; PRT; 89 AA.
 ID O9FH55
 AC O9FH55;
 DT 01-MAR-2001 (TRENBLrel. 16, Created)
 DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
 DT 01-MAR-2001 (TRENBLrel. 16, Last annotation update)
 DE Genomic DNA, chromosome 5, TAC clone KIL20.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 NC NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=COLUMBIA;
 RX MEDLINE=20181125; Pubmed=10718197;
 RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,
 RA Tabata S.,
 RA "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence
 RT features of the regions of 3,076,755 bp covered by sixty P1 and TAC
 RT clones."
 RL DNA Res. 7:31-63 (2000).
 DR EMBL: AB022211; BAB10713.1;
 SQ SEQUENCE 89 AA; 10345 MW; DP12344C836FAE0A CRC64;

Query Match 3.7%; Score 7; DB 10; Length 89;
 Best Local Similarity 100.0%; Pred. No. 26;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 90 RLPLNPI 96
 |||||
 DB 18 RLPLNPI 24

RESULT 10

O9GTU4 PRELIMINARY; PRT; 109 AA.
 ID O9GTU4
 AC O9GTU4;
 DT 01-MAR-2001 (TRENBLrel. 16, Created)
 DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
 DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
 DE Zinc metalloproteinase 3 ME3 (Fragment).
 OS Ancylostoma caninum (Dog hookworm).
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongyloida;
 OC Ancylostomatidae; Ancylostomatidae; Ancylostomatinae; Ancylostoma.
 NC NCBI_TaxID=29170;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Jones B.F., Hotez P.J.,
 RA "Cloning and characterization of a zinc-metalloprotease secreted by
 RT the invasive stages of Ancylostoma caninum."
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF273706; AAG29106.1;
 DR HSSP; P08473; IDMT.
 DR InterPro; IPR000718; Peptidase_M13.
 DR Pfam; PF01431; Peptidase_M13; 1.
 FT NON TER 1
 SQ SEQUENCE 109 AA; 12300 MW; 7C3C54FD724CAD7B CRC64;

Query Match 3.7%; Score 7; DB 5; Length 109;
 Best Local Similarity 100.0%; Pred. No. 30;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 108 DRCIPD 114
 |||||
 DB 86 DRCIPD 92

RESULT 11

O910C0 PRELIMINARY; PRT; 130 AA.
 ID O910C0
 AC O910C0;
 DT 01-MAR-2001 (TRENBLrel. 16, Created)
 DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
 DT 01-OCT-2001 (TRENBLrel. 18, Last annotation update)
 DE Hypothetical protein PA2722.
 GN PA2722.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 NC NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 15692 / PA01;
 RX MEDLINE=20437337; Pubmed=10984043;
 RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
 RA Hickey M.J., Brinkman F.S.L., Hutnagle W.O., Kowalik D.J., Lagrou M.,
 RA Garber R.L., Golety L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
 RA Brady L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 RA Reiter J., Sater M.H., Hancock R.E.W., Lory S., Olson M.V.,
 RA "Complete genome sequence of Pseudomonas aeruginosa PA01, an
 RT opportunistic pathogen."
 RL Nature 406:959-964 (2000).
 DR EMBL: AE004700; AAG06110.1;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 130 AA; 14641 MW; ECEBDCD7F247835E CRC64;

Query Match 3.7%; Score 7; DB 16; Length 130;
 Best Local Similarity 100.0%; Pred. No. 35;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 135 KYRLVAS 141
 |||||
 DB 12 KYRLVAS 18

RESULT 12

O9HVX4 PRELIMINARY; PRT; 148 AA.
 ID O9HVX4
 AC O9HVX4;
 DT 01-MAR-2001 (TRENBLrel. 16, Created)
 DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
 DT 01-OCT-2001 (TRENBLrel. 18, Last annotation update)
 DE Hypothetical protein PA4441.
 GN PA4441.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 NC NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 15692 / PA01;
 RX MEDLINE=20437337; Pubmed=10984043;
 RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
 RA Hickey M.J., Brinkman F.S.L., Hutnagle W.O., Kowalik D.J., Lagrou M.,
 RA Garber R.L., Golety L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
 RA Brady L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 RA Reiter J., Sater M.H., Hancock R.E.W., Lory S., Olson M.V.,
 RA "Complete genome sequence of Pseudomonas aeruginosa PA01, an
 RT opportunistic pathogen."
 RL Nature 406:959-964 (2000).
 DR EMBL: AE004858; AAG07829.1;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 148 AA; 16418 MW; 4DD9E4B217EDD2B8 CRC64;

Query Match 3.7%; Score 7; DB 16; Length 148;
 Best Local Similarity 100.0%; Pred. No. 39;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 89 ARLLPNA 95
 DB 26 ARLLPNA 32

RESULT 13
 ID Q8YQJ5 PRELIMINARY; PRT; 153 AA.

AC Q8YQJ5; 01-MAR-2002 (TREMBlrel. 20, Created)
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Hypothetical protein A13828.

GN ALR3828.
 OS Anabaena sp. (strain PCC 7120).
 OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
 NCBI_TaxId=103690;

RP SEQUENCE FROM N.A.
 RX MEDLINE=21595285; PubMed=11759840; Kuritz T., Sasamoto S.,
 RA Kaneo T., Nakamura Y., Wolk C.P., Kawashima K., Kimura T.,
 RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
 RA Nakazaki N., Shimo S., Sugimoto M., Takazawa M., Yamada M.,
 RA Yasuda M., Tabata S.;

RT "Complete genomic sequence of the filamentous nitrogen-fixing
 cyanobacterium Anabaena sp. strain PCC 7120.";
 RL DNA Res. 8:205-213(2001).

DR EMBL: AP003594; BAB75527.1; -
 DR InterPro: IPR003728; DUF150.

DR Pfam: PF02576; DUF150; 1.
 KW Hypothetical protein; Complete proteome.

SEQUENCE 153 AA; 17259 MW; EED82DIFED43EFC6 CRC64;

Query Match 3.7%; Score 7; DB 16; Length 153;
 Best Local Similarity 100.0%; Pred. No. 40;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 DATEIIP 15
 DB 66 DATEIIP 72

RESULT 14
 ID Q31243 PRELIMINARY; PRT; 183 AA.

AC Q31243; 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE MHC class II A-beta 2 (Fragment).

OS Mus saxicola (Spiry mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

NCBI_TaxId=10094;

RP SEQUENCE FROM N.A.
 RX MEDLINE=90316177; PubMed=2369919;
 RA Cam P., Jouvin-Marche E., Leguern C., Marche P.N.;

RT "Structure of class II genes in wild mouse Mus saxicola: Functional
 and evolutionary implications.";
 RL Eur. J. Immunol. 20:1337-1343(1990).

DR EMBL: M30157; AAA39585.1; -
 DR EMBL: M30156; AAA39585.1; JOINED.

DR HSSP: P13760; 2SER.

DR InterPro: IPR003597; IG_C1.

DR InterPro: IPR003006; IG_MHC.

DR InterPro: IPR000353; MHC_II_beta.

DR Pfam: PF00047; Ig_1.
 DR Pfam: PF00969; MHC_II_beta.1.
 DR ProDom: PD000328; MHC_II_beta.1.
 DR SMART: SM00407; IGc1_1.

DR PROSITE: PS00290; IG_MHC; UNKNOWN_1.
 KW Glycoprotein; MHC II; Transmembrane.

FT NON TER 1
 FT TER 1

SEQUENCE 183 AA; 20906 MW; 79471BF03050AAC9 CRC64;

Query Match 3.7%; Score 7; DB 7; Length 183;
 Best Local Similarity 100.0%; Pred. No. 47;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 EIPPLG 18
 DB 157 EIPPLG 163

RESULT 15
 ID Q9SDR7 PRELIMINARY; PRT; 186 AA.

AC Q9SDR7; 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Dirigent protein.

GN PSD Fil.

OS Forsydia intermedia (Border forsythia).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asterales; euasterids I; Lamiales; Oleaceae; Forsythia.

NCBI_TaxId=55183;

RP SEQUENCE FROM N.A.

RC STRAIN=CV. LYWOOD GOLD; TISSUE=STEM;
 RX MEDLINE=99175375; PubMed=10074466;

RA Gang D.R., Costa M.A., Fujita M., Dinkova-Kostova A.T., Wang H.B.,
 RA Butler V., Martin W., Sarkkanen S., Davin L.B., Lewis N.G.;

RT "Regiochemical control of monolignol radical coupling: a new paradigm
 for lignin and lignan biosynthesis.";

RL Chem. Biol. 6:143-151(1999).

DR EMBL: AF210061; AAF25357.1; -
 DR InterPro: IPR004265; Disease_resp.

DR Pfam: PF03018; disease_resp.1.
 SEQUENCE 186 AA; 20953 MW; 18F73DBCDA4D17F1 CRC64;

Query Match 3.7%; Score 7; DB 10; Length 186;
 Best Local Similarity 100.0%; Pred. No. 47;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 169 GRKPRR 175
 DB 24 GRKPRR 30

Search completed: March 28, 2003, 14:23:22
 Job time : 28.8533 secs

GenCore version 5.1.4.P5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 28, 2003, 14:20:52 ; Search time 11.6533 Seconds
(without alignments)
479.722 Million cell updates/sec

Title: US-09-867-274-2

Perfect score: 190
Sequence: 1 OQMOAFKNDATETIPELGEY.....KRPFRASAKANQAELENAY 190

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 262574 seqs, 29422922 residues

Word size : 0

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

Issued Patents AA: *
1: /cgn2_6/pcodata/1/iaa/5A_COMB.pep: *
2: /cgn2_6/pcodata/1/iaa/5B_COMB.pep: *
3: /cgn2_6/pcodata/1/iaa/5A_COMB.pep: *
4: /cgn2_6/pcodata/1/iaa/5B_COMB.pep: *
5: /cgn2_6/pcodata/1/iaa/PCTUS_COMB.pep: *
6: /cgn2_6/pcodata/1/iaa/Backfill1.pep: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	190	100.0	213	4	US-09-449-218D-2
2	175	92.1	213	4	US-09-449-218D-6
3	175	92.1	213	4	US-09-449-218D-8
4	128	67.4	213	4	US-09-449-218D-10
5	76	40.0	176	4	US-09-449-218D-16
6	33	17.4	213	4	US-09-449-218D-14
7	32	16.8	211	4	US-09-449-218D-12
8	9	4.7	206	1	US-08-468-847B-2
9	9	4.7	206	1	US-08-468-847B-20
10	9	4.7	206	4	US-09-188-930-159
11	9	4.7	206	4	US-09-188-930-159
12	7	3.7	186	4	US-09-475-316A-13
13	7	3.7	917	2	US-08-588-983-16
14	7	3.7	917	2	US-08-588-976-16
15	6	3.2	32	4	US-09-082-358B-70
16	6	3.2	39	3	US-08-554-385-22
17	6	3.2	68	1	US-08-836-791-5
18	6	3.2	80	1	US-08-011-398B-15
19	6	3.2	80	1	US-08-464-051-15
20	6	3.2	80	2	US-08-462-498-15
21	6	3.2	80	3	US-08-554-385-14
22	6	3.2	87	4	US-09-299-495F-31
23	6	3.2	88	2	US-08-690-011A-31
24	6	3.2	97	2	US-08-245-511-30
25	6	3.2	123	4	US-08-600-993A-30
26	6	3.2	123	4	US-09-134-001C-3283
27	6	3.2	129	4	US-09-325-932A-50

28	6	3.2	182	4	US-09-230-670C-6	Sequence 6, Appli
29	6	3.2	186	1	US-07-960-981-5	Sequence 5, Appli
30	6	3.2	186	5	PCT-US93-09634-5	Sequence 5, Appli
31	6	3.2	187	4	US-09-1134-001C-2866	Sequence 2666, Ap
32	6	3.2	199	4	US-09-325-932A-46	Sequence 46, Appli
33	6	3.2	216	2	US-08-821-637-3	Sequence 46, Appli
34	6	3.2	217	4	US-09-134-001C-4620	Sequence 4620, Ap
35	6	3.2	223	4	US-08-928-941D-22	Sequence 22, Appli
36	6	3.2	223	4	US-09-280-590A-22	Sequence 22, Appli
37	6	3.2	224	4	US-09-134-001C-4247	Sequence 4247, Ap
38	6	3.2	229	4	US-09-134-001C-5029	Sequence 5029, Ap
39	6	3.2	243	2	US-08-319-376-2	Sequence 2, Appli
40	6	3.2	248	2	US-08-701-935-3	Sequence 3, Appli
41	6	3.2	248	3	US-09-134-591-3	Sequence 3, Appli
42	6	3.2	254	2	US-08-701-935-6	Sequence 6, Appli
43	6	3.2	254	3	US-09-134-591-6	Sequence 304, Ap
44	6	3.2	255	4	US-09-134-001C-3304	Sequence 3, Appli
45	6	3.2	269	2	US-08-727-311-3	

ALIGNMENTS

```
RESULT 1
US-09-449-218D-2
; Sequence 2, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepfer, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 213
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-449-218D-2

Query Match      100.0%; Score 190; DB 4; Length 213;
Best Local Similarity 100.0%; Pred. No. 8.9e-178;
Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 OQMOAFKNDATETIPELGEYPPPELNNKTMRRAENGSRPHHPFTKQVSEYSCREL 60
DB 24 OQMOAFKNDATETIPELGEYPPPELNNKTMRRAENGSRPHHPFTKQVSEYSCREL 83
QY 61 HFTRTVTGPGCSAKPYTELCSGCGGPALLPNAIGKMMRPSGPPFRICIPRYRAQR 120
DB 84 HFTRTVTGPGCSAKPYTELCSGCGGPALLPNAIGKMMRPSGPPFRICIPRYRAQR 143
QY 121 VOLLCPGGEAPRRKRVRLVASCCKRLTRFNQSELKDFTEARPOKGRKRPARRSAK 180
DB 144 VOLLCPGGEAPRRKRVRLVASCCKRLTRFNQSELKDFTEARPOKGRKRPARRSAK 203
QY 181 ANQAELENAY 190
DB 204 ANQAELENAY 213

RESULT 2
US-09-449-218D-6
; Sequence 6, Application US/09449218D
; Patent No. 6395511
```

```
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepfer, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 213
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-449-218D-6

Query Match          92.1%; Score 175; DB 4; Length 213;
Best Local Similarity 100.0%; Pred. No. 3.8e-163;
Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 ELGEPPEPPLENNKTNRAENGRRPPHPEFTKDVSEYSCRELHTRVYTDGPGCSAK 75
DB 39 ELGEPPEPPLENNKTNRAENGRRPPHPEFTKDVSEYSCRELHTRVYTDGPGCSAK 98
QY 76 PVTSLVSGGCGPARLLPNAIGRGKMWRRSGDPFRCIPDRYRAORVOLLCPGGEAPRARK 135
DB 99 PVTSLVSGGCGPARLLPNAIGRGKMWRRSGDPFRCIPDRYRAORVOLLCPGGEAPRARK 158
QY 136 VRLVASCKCKRLTRFHNSSELDKDFGTAAAPQOKRKRPPRARSKANQAELNNAY 190
DB 159 VRLVASCKCKRLTRFHNSSELDKDFGTAAAPQOKRKRPPRARSKANQAELNNAY 213

RESULT 3
US-09-449-218D-8
; Sequence 8, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepfer, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 213
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-449-218D-8

Query Match          92.1%; Score 175; DB 4; Length 213;
Best Local Similarity 100.0%; Pred. No. 3.8e-163;
Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 ELGEPPEPPLENNKTNRAENGRRPPHPEFTKDVSEYSCRELHTRVYTDGPGCSAK 75
DB 39 ELGEPPEPPLENNKTNRAENGRRPPHPEFTKDVSEYSCRELHTRVYTDGPGCSAK 98
QY 76 PVTSLVSGGCGPARLLPNAIGRGKMWRRSGDPFRCIPDRYRAORVOLLCPGGEAPRARK 135
DB 99 PVTSLVSGGCGPARLLPNAIGRGKMWRRSGDPFRCIPDRYRAORVOLLCPGGEAPRARK 158
QY 136 VRLVASCKCKRLTRFHNSSELDKDFGTAAAPQOKRKRPPRARSKANQAELNNAY 190
DB 159 VRLVASCKCKRLTRFHNSSELDKDFGTAAAPQOKRKRPPRARSKANQAELNNAY 213
```

```
DB 99 PVTSLVSGGCGPARLLPNAIGRGKMWRRSGDPFRCIPDRYRAORVOLLCPGGEAPRARK 158
QY 136 VRLVASCKCKRLTRFHNSSELDKDFGTAAAPQOKRKRPPRARSKANQAELNNAY 190
DB 159 VRLVASCKCKRLTRFHNSSELDKDFGTAAAPQOKRKRPPRARSKANQAELNNAY 213

RESULT 4
US-09-449-218D-10
; Sequence 10, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepfer, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 213
; TYPE: PRT
; ORGANISM: Cercopithecus pygerythrus
US-09-449-218D-10

Query Match          67.4%; Score 128; DB 4; Length 213;
Best Local Similarity 100.0%; Pred. No. 2.7e-117;
Matches 128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGMQAFKNDATETIIPELGEYEPPELNNKTNRAENGRRPPHPEFTKDVSEYSCREL 60
DB 24 QGMQAFKNDATETIIPELGEYEPPELNNKTNRAENGRRPPHPEFTKDVSEYSCREL 83
QY 61 HFTRYVTDGPGCSAKPVTSLVSGGCGPARLLPNAIGRGKMWRRSGDPFRCIPDRYRAOR 120
DB 84 HFTRYVTDGPGCSAKPVTSLVSGGCGPARLLPNAIGRGKMWRRSGDPFRCIPDRYRAOR 143
QY 121 VOLLCPGG 128
DB 144 VOLLCPGG 151

RESULT 5
US-09-449-218D-16
; Sequence 16, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepfer, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 16
; LENGTH: 176
; TYPE: PRT
; ORGANISM: Bos taurus
US-09-449-218D-16
```

Query Match 40.0%; Score 76; DB 4; Length 176;
Best Local Similarity 100.0%; Pred. No. 1.2e-66;
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 53 SEVSCRELHFTRYVTDGPCRSAPVTELVCSGCCGPARLLPNAIGRGKXWRPSGDPFRCT 112
DB 45 SEVSCRELHFTRYVTDGPCRSAPVTELVCSGCCGPARLLPNAIGRGKXWRPSGDPFRCT 104

QY 113 PDRYRQRYVQLCPGG 128
DB 105 PDRYRQRYVQLCPGG 120

RESULT 6
US-09-449-218D-14

; Sequence 14, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepfer, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 213
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-449-218D-14

Query Match 17.4%; Score 33; DB 4; Length 213;
Best Local Similarity 100.0%; Pred. No. 1.3e-24;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 66 VTGDCRSAPVTELVCSGCCGPARLLPNAIGR 98
DB 89 VTGDCRSAPVTELVCSGCCGPARLLPNAIGR 121

RESULT 7
US-09-449-218D-12

; Sequence 12, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepfer, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 12
; LENGTH: 211
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-449-218D-12

Query Match 16.8%; Score 32; DB 4; Length 211;
Best Local Similarity 100.0%; Pred. No. 1.2e-23;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 TDGPCRSAPVTELVCSGCCGPARLLPNAIGR 98
DB 88 TDGPCRSAPVTELVCSGCCGPARLLPNAIGR 119

RESULT 8
US-08-468-847B-2

; Sequence 2, Application US/08468847B
; Patent No. 5780263
; GENERAL INFORMATION:
; APPLICANT: Hastings, Gregg A. and Adams, Mark D.
; TITLE OF INVENTION: Human CCN-Like Growth Factor
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLIAN,
; ADDRESSEE: CECCHI, STEWART & OLSTEIN
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07068

; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,847B
; FILING DATE: 6 June 1995
; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:

; ATTORNEY/AGENT INFORMATION:
; NAME: MULHINS, J.G.
; REGISTRATION NUMBER: 33,073

; REFERENCE/DOCKET NUMBER: 325800-442
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744

; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 206 AMINO ACIDS
; TYPE: AMINO ACID
; STRANDEDNESS:

; TOPOLOGY: LINEAR
; MOLECULE TYPE: PROTEIN
US-08-468-847B-2

Query Match 4.7%; Score 9; DB 1; Length 206;
Best Local Similarity 100.0%; Pred. No. 0.32;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 AFKNDATETI 13
DB 23 AFKNDATETI 31

RESULT 9
US-08-468-847B-20

; Sequence 20, Application US/08468847B
; Patent No. 5780263
; GENERAL INFORMATION:
; APPLICANT: Hastings, Gregg A. and Adams, Mark D.
; TITLE OF INVENTION: Human CCN-Like Growth Factor
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLIAN,
; ADDRESSEE: CECCHI, STEWART & OLSTEIN

```

STREET: 6 BECKER FARM ROAD
CITY: ROSELAND
STATE: NEW JERSEY
COUNTRY: USA
ZIP: 07068
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 INCH DISKETTE
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WORD PERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/468,847B
FILING DATE: 6 June 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: MULLINS, J.G.
REGISTRATION NUMBER: 33,073
REFERENCE/DOCKET NUMBER: 325800-442
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 206 AMINO ACIDS
TYPE: AMINO ACID
STRANDEDNESS:
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
US-08-468-847B-20

```

```

Query Match          4.7%; Score 9; DB 1; Length 206;
Best Local Similarity 100.0%; Pred. No. 0.32;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 5 AFKNDATETI 13
    |||||
Db 23 AFKNDATETI 31

```

```

RESULT 10
US-09-188-930-159
; Sequence 159, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James Greg
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; TITLE OF INVENTION: and Methods For Their Use
; FILE REFERENCE: 11000.1011c1
; CURRENT APPLICATION NUMBER: US/09/188,930A
; CURRENT FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 159
; LENGTH: 206
; TYPE: PRT
; ORGANISM: mouse
US-09-188-930-159

```

```

Query Match          4.7%; Score 9; DB 4; Length 206;
Best Local Similarity 100.0%; Pred. No. 0.32;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 5 AFKNDATETI 13
    |||||
Db 23 AFKNDATETI 31

```

```

RESULT 11
US-09-188-930-286
; Sequence 286, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James Greg
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; TITLE OF INVENTION: and Methods For Their Use
; FILE REFERENCE: 11000.1011c1
; CURRENT APPLICATION NUMBER: US/09/188,930A
; CURRENT FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 286
; LENGTH: 206
; TYPE: PRT
; ORGANISM: Mouse
US-09-188-930-286

```

```

Query Match          4.7%; Score 9; DB 4; Length 206;
Best Local Similarity 100.0%; Pred. No. 0.32;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 5 AFKNDATETI 13
    |||||
Db 23 AFKNDATETI 31

```

```

RESULT 12
US-09-475-316A-13
; Sequence 13, Application US/09475316A
; Patent No. 6210942
; GENERAL INFORMATION:
; APPLICANT: Lewis, No. 6210942man G.
; APPLICANT: Davin, Laurence B.
; APPLICANT: Dinkova-Kostova, Albena T.
; APPLICANT: Fujita, Masayuki
; APPLICANT: Gang, David R.
; APPLICANT: Sarkanen, Simo
; APPLICANT: Ford, Joshua D
; TITLE OF INVENTION: RECOMBINANT PINORESINOL/LARICRESINOL REDUCTASES,
; TITLE OF INVENTION: RECOMBINANT DIRIGENT PROTEINS AND METHODS OF USE
; FILE REFERENCE: WSUR-1-13793
; CURRENT APPLICATION NUMBER: US/09/475,316A
; CURRENT FILING DATE: 1999-12-30
; PRIOR APPLICATION NUMBER: 09/307,653
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: PCT/US97/20391
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: 60/054,380
; PRIOR FILING DATE: 1997-07-31
; PRIOR APPLICATION NUMBER: 60/030,522
; PRIOR FILING DATE: 1996-11-08
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 186
; TYPE: PRT
; ORGANISM: Forsythia x intermedia
US-09-475-316A-13

```

```

Query Match          3.7%; Score 7; DB 4; Length 186;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 169 GRKRRPR 175
    |||||
Db 24 GRKRRPR 30

```

RESULT 13
US-08-588-983-16
Sequence 16, Application US/08588983
Patent No. 5854067
GENERAL INFORMATION:
APPLICANT: Christopher B. Newgard, et al.
TITLE OF INVENTION: Methods and Compositions
TITLE OF INVENTION: for Inhibiting Hexokinase
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: TX
COUNTRY: US
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/588,983
FILING DATE: Concurrently herewith
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Fussey, Shelley P.M.
REGISTRATION NUMBER: 39,458
REFERENCE/DOCKET NUMBER: UTSD:424/FUS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (512) 474-7577
TELEX: n/a
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 917 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-588-983-16
Query Match
Best Local Similarity 3.7%; Score 7; DB 2; Length 917;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 47 FETKQVS 53
Db 334 FETKQVS 340
RESULT 14
US-08-588-976-16
Sequence 16, Application US/08588976
Patent No. 5891717
GENERAL INFORMATION:
APPLICANT: Christopher B. Newgard, et al.
TITLE OF INVENTION: Methods and Compositions for
TITLE OF INVENTION: Inhibiting Hexokinase
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: TX
COUNTRY: US
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/588,976
FILING DATE: Concurrently herewith
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Fussey, Shelley P.M.
REGISTRATION NUMBER: 39,458
REFERENCE/DOCKET NUMBER: UTSD:481/FUS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (512) 474-7577
TELEX: n/a
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 917 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-588-976-16

Query Match
Best Local Similarity 3.7%; Score 7; DB 2; Length 917;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 47 FETKQVS 53
Db 334 FETKQVS 340

RESULT 15
US-09-082-358B-70
Sequence 70, Application US/09082358B
Patent No. 6469153
GENERAL INFORMATION:
APPLICANT: Goff, Stephen P.
TITLE OF INVENTION: EIP-1, EIP-3 GENES, ENVELOPE-INTERACTING PROTEINS,
TITLE OF INVENTION: EIP-1, and EIP-3
FILE REFERENCE: 0575/54804
CURRENT APPLICATION NUMBER: US/09/082,358B
CURRENT FILING DATE: 1998-05-20
NUMBER OF SEQ ID NOS: 106
SOFTWARE: Patentln Ver. 2.1
SEQ ID NO 70
LENGTH: 32
TYPE: PRT
ORGANISM: murine
US-09-082-358B-70

Query Match
Best Local Similarity 3.2%; Score 6; DB 4; Length 32;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 128 GEAPRA 133
Db 10 GEAPRA 15

Search completed: March 28, 2003, 14:24:27
Job time : 13.6533 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 28, 2003, 14:22:02 ; Search time 12.16 Seconds
(without alignments)
917.557 Million cell updates/sec

Title: US-09-867-274-2

Perfect score: 190
Sequence: 1 OQWQAFKQDATEIIPELGEY.....KRPFRASAKANQAELENAY 190

Scoring table: OLIGO
Gapex 60.0 , Gapext 60.0

Searched: 237916 seqs, 58723674 residues

Word size : 0

Total number of hits satisfying chosen parameters: 237916

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

Published Applications AA.*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	190	100.0	190	US-09-867-274-2	Sequence 2, Appl
2	190	100.0	213	US-09-867-274-5	Sequence 5, Appl
3	139	73.2	139	US-09-864-761-47109	Sequence 47109, A
4	32	16.8	185	US-09-867-274-4	Sequence 4, Appl
5	32	16.8	208	US-09-867-274-6	Sequence 6, Appl
6	9	4.7	206	US-09-866-050A-159	Sequence 159, App
7	9	4.7	206	US-09-866-050A-286	Sequence 286, App
8	9	4.7	206	US-09-853-625B-2	Sequence 2, Appl
9	9	4.7	206	US-09-853-625B-10	Sequence 20, Appl
10	8	4.2	183	US-09-867-274-25	Sequence 25, Appl
11	7	3.7	186	US-09-944-160-16	Sequence 16, Appl
12	7	3.7	205	US-09-864-761-36244	Sequence 36244, A
13	7	3.7	373	US-09-738-626-3658	Sequence 3658, Ap
14	7	3.7	705	US-10-154-386-2	Sequence 2, Appl
15	7	3.7	917	US-09-808-743-10	Sequence 10, Appl
16	6	3.2	46	US-09-925-297-836	Sequence 836, App
17	6	3.2	53	US-10-001-857-164	Sequence 164, App
18	6	3.2	59	US-09-510-332-72	Sequence 72, Appl
19	6	3.2	59	US-09-864-761-34799	Sequence 34799, A

20	6	3.2	87	9	US-10-059-720-31	Sequence 31, Appl
21	6	3.2	87	9	US-10-091-572-334	Sequence 334, App
22	6	3.2	104	10	US-09-864-761-33381	Sequence 33381, A
23	6	3.2	110	10	US-09-741-669-320	Sequence 320, App
24	6	3.2	110	10	US-09-912-020-387	Sequence 387, App
25	6	3.2	110	10	US-09-815-242-10354	Sequence 10354, A
26	6	3.2	110	10	US-09-815-242-13970	Sequence 13970, A
27	6	3.2	112	10	US-09-815-242-5235	Sequence 5235, App
28	6	3.2	113	9	US-10-013-379-17	Sequence 17, Appl
29	6	3.2	117	10	US-09-815-242-12258	Sequence 12258, A
30	6	3.2	117	10	US-09-815-242-12765	Sequence 12765, A
31	6	3.2	124	9	US-09-738-626-5430	Sequence 5430, App
32	6	3.2	124	10	US-09-912-020-250	Sequence 250, App
33	6	3.2	124	10	US-09-815-242-10365	Sequence 10365, A
34	6	3.2	124	10	US-09-815-242-13967	Sequence 13967, A
35	6	3.2	132	9	US-09-981-876-147	Sequence 147, App
36	6	3.2	132	9	US-09-148-545-147	Sequence 147, App
37	6	3.2	137	10	US-09-867-550-832	Sequence 832, App
38	6	3.2	143	10	US-09-815-242-11669	Sequence 11669, A
39	6	3.2	149	10	US-09-764-864-1171	Sequence 1171, App
40	6	3.2	160	10	US-09-815-242-11664	Sequence 11664, A
41	6	3.2	163	9	US-09-925-299-1035	Sequence 1035, App
42	6	3.2	163	10	US-09-925-299-1035	Sequence 1035, App
43	6	3.2	165	9	US-09-738-626-5314	Sequence 5314, App
44	6	3.2	195	10	US-09-858-664A-11	Sequence 11, Appl
45	6	3.2	201	10	US-09-925-300-1646	Sequence 1646, App

ALIGNMENTS

RESULT 1
US-09-867-274-2
Sequence 2, Application US/09867274
Patent No. US20020106650A1
GENERAL INFORMATION:
APPLICANT: Paszty, Christopher
APPLICANT: Gao, Yongming
TITLE OF INVENTION: Cysteine Knot Polypeptides: Cloaked-2 Molecules and Uses Thereof
FILE REFERENCE: 01017/37428
CURRENT APPLICATION NUMBER: US/09/867,274
CURRENT FILING DATE: 2001-05-29
PRIOR APPLICATION NUMBER: US 60/208,550
PRIOR FILING DATE: 2000-06-01
PRIOR APPLICATION NUMBER: US 60/223,542
PRIOR FILING DATE: 2000-08-04
NUMBER OF SEQ ID NOS: 25
SOFTWARE: PatentIn version 3.0
SEQ ID NO 2
LENGTH: 190
TYPE: PRT
ORGANISM: Homo sapiens
US-09-867-274-2

Query Match 100.0%; Score 190; DB 10; Length 190;
Best Local Similarity 100.0%; Pred. No. 2.5e-180;
Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 OQWQAFKQDATEIIPELGEYPPPELENNKTNRAENGSRPHHPFTKQVSEYSCREL 60
DB 1 OQWQAFKQDATEIIPELGEYPPPELENNKTNRAENGSRPHHPFTKQVSEYSCREL 60
QY 61 HFTFRVVTGPGCSAKPVELVCSGCCGPARLLPNAIGGKMWBPSPGDFRCIPRYRAQR 120
DB 61 HFTFRVVTGPGCSAKPVELVCSGCCGPARLLPNAIGGKMWBPSPGDFRCIPRYRAQR 120
QY 121 VOLLCPGSGAPRAKRVIRVASCCKRLTRFNQSELKDFGTEDARPQGRKPRPARSAK 180
DB 121 VOLLCPGSGAPRAKRVIRVASCCKRLTRFNQSELKDFGTEDARPQGRKPRPARSAK 180
QY 181 ANQAELENAY 190
DB 181 ANQAELENAY 190

```
RESULT 2
US-09-867-274-5
; Sequence 5, Application US/09867274
; Patent No. US20020106650A1
; GENERAL INFORMATION:
; APPLICANT: Paszty, Christopher
; APPLICANT: Gao, Yongming
; TITLE OF INVENTION: Cysteine Knot Polypeptides: Cloaked-2 Molecules and Uses Thereof
; FILE REFERENCE: 01017/37428
; CURRENT APPLICATION NUMBER: US/09/867,274
; PRIOR FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: US 60/208,550
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: US 60/223,542
; PRIOR FILING DATE: 2000-08-04
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 213
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-867-274-5

Query Match      100.0%; Score 190; DB 10; Length 213;
Best Local Similarity 100.0%; Pred. No. 2.7e-180;
Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGMQAFKNDATETIPELGEYEPPEPELENNKTKMRAENGSRPHHPETDVSYSREL 60
DB 24 QGMQAFKNDATETIPELGEYEPPEPELENNKTKMRAENGSRPHHPETDVSYSREL 83
QY 61 HFTRYVTDDPCRSKAPVTETVCSGCCGPARLLPNAIGRGKMWPRSGDPFCIPDRTYRQR 120
DB 84 HFTRYVTDDPCRSKAPVTETVCSGCCGPARLLPNAIGRGKMWPRSGDPFCIPDRTYRQR 143
QY 121 VOLLCGGEAPRARKVRLVASCCKRLTRFHNOSLKDQFGTEARPOCKGRKPRPARSAK 180
DB 144 VOLLCGGEAPRARKVRLVASCCKRLTRFHNOSLKDQFGTEARPOCKGRKPRPARSAK 203
QY 181 ANQAELENAY 190
DB 204 ANQAELENAY 213

RESULT 3
US-09-864-761-47109
; Sequence 47109, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
```

```
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 47109
; LENGTH: 139
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC003098.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.62
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.6
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.49
; OTHER INFORMATION: SWISSPROT HIT: P45646, EVALU 4.70e-01
US-09-864-761-47109

Query Match      73.2%; Score 139; DB 10; Length 139;
Best Local Similarity 100.0%; Pred. No. 5.1e-130;
Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 52 VSEYSGRELHFTRYVTDDPCRSKAPVTETVCSGCCGPARLLPNAIGRGKMWPRSGDPFC 111
DB 1 VSEYSGRELHFTRYVTDDPCRSKAPVTETVCSGCCGPARLLPNAIGRGKMWPRSGDPFC 60
QY 112 IPDRYRAGRVLCLCGGEAPRARKVRLVASCCKRLTRFHNOSLKDQFGTEARPOCKGRK 171
DB 61 IPDRYRAGRVLCLCGGEAPRARKVRLVASCCKRLTRFHNOSLKDQFGTEARPOCKGRK 120
QY 172 PRPARSAKANQAELENAY 190
DB 121 PRPARSAKANQAELENAY 139

RESULT 4
US-09-867-274-4
; Sequence 4, Application US/09867274
; Patent No. US20020106650A1
; GENERAL INFORMATION:
; APPLICANT: Paszty, Christopher
; APPLICANT: Gao, Yongming
; TITLE OF INVENTION: Cysteine Knot Polypeptides: Cloaked-2 Molecules and Uses Thereof
; FILE REFERENCE: 01017/37428
; CURRENT APPLICATION NUMBER: US/09/867,274
; PRIOR FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: US 60/208,550
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: US 60/223,542
; PRIOR FILING DATE: 2000-08-04
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 185
```

TYPE: PRT
ORGANISM: Mus musculus
US-09-867-274-4

Query Match 16.8%; Score 32; DB 10; Length 185;
Best Local Similarity 100.0%; Pred. No. 4,4e-24;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 67 TDGPCRSAPVTELVCSGCCPARLLPNAIGR 98
Db 65 TDGPCRSAPVTELVCSGCCPARLLPNAIGR 96

RESULT 5
US-09-867-274-6
Sequence 6, Application US/09867274
Patent No. US20020106650A1
GENERAL INFORMATION:
APPLICANT: Paszty, Christopher
APPLICANT: Gao, Yongming
TITLE OF INVENTION: Cysteine Knot Polypeptides: Cloaked-2 Molecules and Uses Thereof
FILE REFERENCE: 01017/37428
CURRENT APPLICATION NUMBER: US/09/867,274
CURRENT FILING DATE: 2001-05-29
PRIOR APPLICATION NUMBER: US 60/208,550
PRIOR FILING DATE: 2000-06-01
PRIOR APPLICATION NUMBER: US 60/223,542
PRIOR FILING DATE: 2000-08-04
NUMBER OF SEQ ID NOS: 25
SOFTWARE: PatentIn version 3.0
SEQ ID NO 6
LENGTH: 208
TYPE: PRT
ORGANISM: Mus musculus
US-09-867-274-6

Query Match 16.8%; Score 32; DB 10; Length 208;
Best Local Similarity 100.0%; Pred. No. 4.9e-24;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 67 TDGPCRSAPVTELVCSGCCPARLLPNAIGR 98
Db 88 TDGPCRSAPVTELVCSGCCPARLLPNAIGR 119

RESULT 6
US-09-866-050A-159
Sequence 159, Application US/09866050A
Publication No. US20030040471A1
GENERAL INFORMATION:
APPLICANT: Watson, James D.
APPLICANT: Strachan, Lorna
APPLICANT: Sleeman, Matthew
APPLICANT: Onrust, Rene
APPLICANT: Murison, James G.
APPLICANT: Kumble, Krishanand D.
TITLE OF INVENTION: Compositions Isolated From Skin Cells
FILE REFERENCE: 11000.1011c4U
CURRENT APPLICATION NUMBER: US/09/866,050A
CURRENT FILING DATE: 2001-05-24
NUMBER OF SEQ ID NOS: 725
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 159
LENGTH: 206
TYPE: PRT
ORGANISM: Mouse
US-09-866-050A-159

Query Match 4.7%; Score 9; DB 9; Length 206;
Best Local Similarity 100.0%; Pred. No. 0.27;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 AFKNDATERI 13
Db 23 AFKNDATERI 31

RESULT 7
US-09-866-050A-286
Sequence 286, Application US/09866050A
Publication No. US20030040471A1
GENERAL INFORMATION:
APPLICANT: Watson, James D.
APPLICANT: Strachan, Lorna
APPLICANT: Sleeman, Matthew
APPLICANT: Onrust, Rene
APPLICANT: Murison, James G.
APPLICANT: Kumble, Krishanand D.
TITLE OF INVENTION: Compositions Isolated From Skin Cells
FILE REFERENCE: 11000.1011c4U
CURRENT APPLICATION NUMBER: US/09/866,050A
CURRENT FILING DATE: 2001-05-24
NUMBER OF SEQ ID NOS: 725
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 286
LENGTH: 206
TYPE: PRT
ORGANISM: Mouse
US-09-866-050A-286

Query Match 4.7%; Score 9; DB 9; Length 206;
Best Local Similarity 100.0%; Pred. No. 0.27;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 AFKNDATERI 13
Db 23 AFKNDATERI 31

RESULT 8
US-09-853-625B-2
Sequence 2, Application US/09853625B
Patent No. US20020049304A1
GENERAL INFORMATION:
APPLICANT: Hastings, Gregg A. and Adams, Mark D.
TITLE OF INVENTION: Human CCN-Like Growth Factor
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLIAN,
CECCHI, STEWART & OLSTEIN
STREET: 6 BECKER FARM ROAD
CITY: ROSELAND
STATE: NEW JERSEY
COUNTRY: USA
ZIP: 07068
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 INCH DISKETTE
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WORD PERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/853,625B
FILING DATE: 14-May-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/053,587
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: MULLINS, J.G.
REGISTRATION NUMBER: 33,073
REFERENCE/DOCKET NUMBER: 325800-442
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 206 AMINO ACIDS
TYPE: AMINO ACID
STRANDEDNESS: <Unknown>
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-853-625B-2

Query Match 4.7%; Score 9; DB 10; Length 206;
Best Local Similarity 100.0%; Pred. No. 0.27;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 AFKXDATEI 13
DB 23 AFKXDATEI 31

RESULT 9
US-09-853-625B-20
Sequence 20, Application US/09853625B
Patent No. US20020049304A1
GENERAL INFORMATION:
APPLICANT: Hastings, Gregg A. and Adams, Mark D.
TITLE OF INVENTION: Human CCN-Like Growth Factor
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESSES:
ADDRESS: CARELIA, BYRNE, BAIN, GILFILLIAN,
CECCHI, STEWART & OLSTEIN
STREET: 6 BECKER FARM ROAD
CITY: ROSELAND
STATE: NEW JERSEY
COUNTRY: USA
ZIP: 07068
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 INCH DISKETTE
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WORD PERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/853,625B
FILING DATE: 14-May-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/053,587
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: MULINS, J.G.
REGISTRATION NUMBER: 33,073
REFERENCE/DOCKET NUMBER: 325800-442
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 206 AMINO ACIDS
TYPE: AMINO ACID
STRANDEDNESS: <Unknown>
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
SEQUENCE DESCRIPTION: SEQ ID NO: 20:
US-09-853-625B-20

Query Match 4.7%; Score 9; DB 10; Length 206;
Best Local Similarity 100.0%; Pred. No. 0.27;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 AFKXDATEI 13
DB 23 AFKXDATEI 31

RESULT 10
US-09-867-274-25
Sequence 25, Application US/09867274
Patent No. US20020106650A1
GENERAL INFORMATION:
APPLICANT: Paszty, Christopher
TITLE OF INVENTION: Cysteine Knot Polypeptides: Cloaked-2 Molecules and Uses Thereof
FILE REFERENCE: 01017/37428
CURRENT APPLICATION NUMBER: US/09/867,274
PRIOR FILING DATE: 2001-05-29
PRIOR APPLICATION NUMBER: US 60/208,550
PRIOR FILING DATE: 2000-06-01
PRIOR APPLICATION NUMBER: US 60/223,542
PRIOR FILING DATE: 2000-08-04
NUMBER OF SEQ ID NOS: 25
SOFTWARE: PatentIn version 3.0
SEQ ID NO 25
LENGTH: 183
TYPE: PRT
ORGANISM: Homo sapiens
US-09-867-274-25

Query Match 4.2%; Score 8; DB 10; Length 183;
Best Local Similarity 100.0%; Pred. No. 2.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 FKXDATEI 13
DB 1 FKXDATEI 8

RESULT 11
US-09-944-160-16
Sequence 16, Application US/09944160
Patent No. US20020174452A1
GENERAL INFORMATION:
APPLICANT: Lewis, No. US20020174452A1man
APPLICANT: Davin, Laurence
TITLE OF INVENTION: Monocot Seeds with Increased Lignin
FILE REFERENCE: WSUR117983
CURRENT APPLICATION NUMBER: US/09/944,160
CURRENT FILING DATE: 2001-08-30
PRIOR APPLICATION NUMBER: US 60/230,632
PRIOR FILING DATE: 2000-09-07
NUMBER OF SEQ ID NOS: 49
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 16
LENGTH: 186
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: dirigent amino acid sequence from plasmid pAP1244
US-09-944-160-16

Query Match 3.7%; Score 7; DB 9; Length 186;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 169 GRKPRR 175
DB 24 GRKPRR 30

RESULT 12
US-09-864-761-36244
Sequence 36244, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.

APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aemica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263,6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 36244
LENGTH: 205
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AL050305.9
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.93
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.99
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.99
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.92
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.9
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.6
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.3
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.96
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.9
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2
OTHER INFORMATION: EST HUMAN HIT: A181773.1, EVALUATE 1.00e-17
OTHER INFORMATION: SWISSPROT HIT: Q09472, EVALUATE 2.40e-01
US-09-864-761-36244

Query Match 3.7%; Score 7; DB 10; Length 205;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 11 TRITPEL 17
DB 145 TRITPEL 151

RESULT 13
US-09-738-626-3568
Sequence 3568, Application US/09738626
Publication No. US20020197605A1
GENERAL INFORMATION:
APPLICANT: NAKAGAWA, SATOSHI
APPLICANT: MIZOGUCHI, HIROSHI
APPLICANT: ANDO, SEIKO
APPLICANT: HAYASHI, MIKIRO
APPLICANT: OCHIAI, KEIKO
APPLICANT: YOKOI, HARUHIKO
APPLICANT: TATEISHI, NAOKO
APPLICANT: SENOH, AKIHIRO
APPLICANT: IKEDA, MASATO
APPLICANT: OZAKI, AKIO
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738,626
CURRENT FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: JP 99/377484
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: PatentIn ver. 3.0
SEQ ID NO 3568
LENGTH: 373
TYPE: PRT
ORGANISM: Corynebacterium glutamicum
US-09-738-626-3568

Query Match 3.7%; Score 7; DB 9; Length 373;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 63 TRYVTDG 69
DB 39 TRYVTDG 45

RESULT 14
US-10-154-386-2
Sequence 2, Application US/10154386
Publication No. US20030026793A1
GENERAL INFORMATION:
APPLICANT: Angiogene Inc.
TITLE OF INVENTION: HIPPOXIA INDUCING FACTORS AND USES THEREOF FOR INDUCING ANGIOGENESIS
FILE REFERENCE: 5600-81
CURRENT APPLICATION NUMBER: US/10/154,386
CURRENT FILING DATE: 2002-05-23
PRIOR APPLICATION NUMBER: US 60/292,630
PRIOR FILING DATE: 2001-05-22
PRIOR APPLICATION NUMBER: US 60/354529
PRIOR FILING DATE: 2002-02-08
NUMBER OF SEQ ID NOS: 13
SOFTWARE: PatentIn version 3.1
SEQ ID NO 2
LENGTH: 705
TYPE: PRT
ORGANISM: Homo sapiens
US-10-154-386-2

Query Match 3.7%; Score 7; DB 9; Length 705;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 172 PRPARS 178

Db 556 PRPRARS 562

RESULT 15

US-09-808-743-10
; Sequence 10, Application US/09808743
; Patent No. US20020068711A1
; GENERAL INFORMATION:
; APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE
; APPLICANT: PEDERSEN, Peter
; APPLICANT: MATHUPALA, Saroj
; TITLE OF INVENTION: ARREST OF PROLIFERATION OF HIGHLY GLYCOLYTIC TUMORS
; FILE REFERENCE: JHU1720-1
; CURRENT APPLICATION NUMBER: US/09/808,743
; CURRENT FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: US 60/189,222
; PRIOR FILING DATE: 2000-03-14
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 10
; LENGTH: 917
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-808-743-10

Query Match 3.7%; Score 7; DB 10; Length 917;

Best local Similarity 100.0%; Pred. No. 97;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 FETKQVS 53
|||||

Db 334 FETKQVS 340

Search completed: March 28, 2003, 14:24:57
Job time : 14.16 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 28, 2003, 17:27:08 ; Search time 1369.11 Seconds
(without alignments)
7523.378 Million cell updates/sec

Title: US-09-867-274-3

Perfect score: 636
Sequence: 1 atgcagccctcactagcccc.....agctggagaa-gcctactag 636

Scoring table: IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_estbm:*
3: em_estbn:*
4: em_estbu:*
5: em_estbv:*
6: em_estbl:*
7: em_estro:*
8: em_estr:*
9: gb_est1:*
10: gb_est2:*
11: gb_est3:*
12: gb_est4:*
13: gb_est5:*
14: gb_est6:*
15: em_estfun:*
16: em_estom:*
17: gb_gsa:*
18: em_gsa_hum:*
19: em_gsa_inv:*
20: em_gsa_pin:*
21: em_gsa_vrt:*
22: em_gsa_fun:*
23: em_gsa_mam:*
24: em_gsa_mus:*
25: em_gsa_other:*
26: em_gsa_pro:*
27: em_gsa_tod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	636	100.0	1990	11 AK017295	AK017295 Mus muscu
2	608	95.6	627	10 BB636457	BB636457 Mus muscu
3	583	91.7	663	10 BB637315	BB637315 BB637315
4	469.4	73.8	511	10 BB638050	BB638050 BB638050
5	391.4	61.5	419	10 BE101082	BE101082 UI-R-B11-
6	336.6	52.9	360	9 A1556282	A1556282 UI-R-C2p-

C	7	289.4	45.5	315	9	A1113131	A1113131 UI-R-C2p-
C	8	273.4	43.0	291	10	BE111224	BE111224 UI-R-B11-
C	9	138.8	21.8	535	17	AQ171546	AQ171546 HS_3088_B
10	136.2	21.4	254	12	BF523030	BF523030 UI-R-C2p-	
11	83.2	13.1	1612	11	AK002396	AK002396 Mus muscu	
12	83.2	13.1	1613	11	AK007893	AK007893 Mus muscu	
13	83.2	13.1	1690	11	AK002240	AK002240 Mus muscu	
14	83.2	13.1	1691	11	AK007967	AK007967 Mus muscu	
15	82.2	12.9	669	12	BF607657	BF607657 MYL_00059	
16	80.4	12.6	472	12	BF420024	BF420024 UI-R-B12-	
17	80.4	12.6	532	12	BF407514	BF407514 UI-R-B12-	
18	79.4	12.5	385	10	AM434705	AM434705 UI-R-B12-	
19	78.2	12.3	789	11	AK007935	AK007935 Mus muscu	
20	76.4	12.0	687	13	BI081999	BI081999 602879836	
21	76	11.9	623	13	BI143761	BI143761 602907224	
22	74.4	11.7	577	13	BI682960	BI682960 464177 MA	
23	74.2	11.7	584	12	BF041721	BF041721 BP250005B	
24	74.2	11.7	925	17	CNS0091P	AL053013 Drosoph11	
25	73.4	11.5	355	9	AA197904	AA197904 mv05c05.x	
26	73.4	11.5	433	9	AA067619	AA067619 mm23e07.x	
27	73	11.5	320	9	AA605771	AA605771 fa19b09.s	
28	71.4	11.2	925	17	CNS0091P	AL053013 Drosoph11	
29	70.6	11.1	627	17	A2876400	A2876400 2M0191D21	
30	68.4	10.8	683	12	BG705527	BG705527 602685416	
31	68.2	10.7	1020	13	BI767517	BI767517 603061406	
32	68	10.7	932	17	CNS00720	AL066742 Drosoph11	
33	67.2	10.6	715	13	BI489720	BI489720 603032117	
34	67.2	10.6	735	13	BI763021	BI763021 603047812	
35	67.2	10.6	740	13	BI760787	BI760787 603043808	
36	67.2	10.6	850	9	AL524992	AL524992 AL524992	
37	67.2	10.6	1028	14	BQ067667	BQ067667 AGENCOURT	
38	67.2	10.6	1094	14	BM924116	BM924116 AGENCOURT	
39	67.2	10.6	1111	13	BM548167	BM548167 AGENCOURT	
40	67.2	10.6	1873	11	AF361494	AF361494 Homo_sapi	
41	66.4	10.4	597	14	BM751691	BM751691 K-EST0027	
42	65.2	10.3	1076	14	BM921437	BM921437 AGENCOURT	
43	64.8	10.2	578	14	BM709058	BM709058 UI-E-CQ1-	
44	64.8	10.2	618	14	BM748765	BM748765 K-EST0023	
45	64.8	10.2	645	14	BM835363	BM835363 K-EST0110	

ALIGNMENTS

RESULT 1	AK017295	1990 bp	mRNA	linear	HTC 19-JAN-2002
LOCUS	AK017295				
DEFINITION	Mus musculus 6 days neonate head cDNA, RIKEN full-length enriched library, clone:5430411E23:sclerostin, full insert sequence.				
ACCESSION	AK017295				
VERSION	AK017295.1 GI:12856464				
KEYWORDS	HTC; CAP trapper.				
SOURCE	Mus musculus (Strain:C57BL/6J) 6 days neonate head cDNA to mRNA, clone:lib:RIKEN full-length enriched mouse cDNA library				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.				
AUTHORS	Carninci, P. and Hayashizaki, Y.				
TITLE	High-efficiency full-length cDNA cloning				
JOURNAL	Mech. Enzymol. 303, 19-44 (1999)				
MEDLINE	99279253				
PUBMED	10349636				
REFERENCE	2				
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.				
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes				
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)				
MEDLINE	20499374				
PUBMED	11042159				
REFERENCE	3				

AUTHORS

Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P., Komano, H., Akiyama, J., Nishi, K., Katsunari, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

TITLE

RIKEN integrated sequence analysis (RISA) system-384 format

JOURNAL

Genome Res. 10 (11), 1757-1771 (2000)

MEDLINE

20530913

PUBMED

11076861

REFERENCE

AUTHORS

Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Aizawa, K., Hara, A., Fukunishi, Y., Komano, H., Adachi, T., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuenl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G., Quackenbush, J., Schriml, L.M., Staudl, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Maehima, D., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schombach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Togo, O.K., Wang, K.H., Weitz, C., Whitaker, C., Wilming, J., Wyshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohlschki, S. and Hayashizaki, Y.

Functional annotation of a full-length mouse cDNA collection

JOURNAL

Nature 409 (6821), 685-690 (2001)

MEDLINE

21085660

PUBMED

11217851

REFERENCE

AUTHORS

5 (bases 1 to 1990)
Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A., Arakawa, T., Baldarelli, R., Bono, H., Brownstein, M., Bult, C., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hiroaka, T., Horii, F., Hume, D., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Komano, H., Kouda, M., Koyama, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numata, K., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Quackenbush, J., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Schriml, L., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Teijima, Y., Toya, T., Yamamura, T., Yamanka, I., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.

TITLE

Direct Submission

JOURNAL

Submitted (10-JUL-2000)

MEDLINE

Yoshida, K.

PUBMED

11217851

REFERENCE

AUTHORS

COMMENT

Yoshida, K.

TITLE

Submitted (10-JUL-2000)

JOURNAL

Yoshida, K.

MEDLINE

11217851

PUBMED

11217851

REFERENCE

AUTHORS

Yoshida, K.

FEATURES

source

1. 1990

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="PANTOM DB:5430411E23"

/db_xref="MGI:1907679"

/db_xref="taxon:10090"

/clone="5430411E23"

/issue_type="head"

/clone_lib="RIKEN full-length enriched mouse cDNA library"

/dev_stage="6 days neonate"

1. 1990

/gene="Soat"

/gene="Soat"

/note="data source:MGI, source key:MGI.1921749, evidence:ISS"

putative

sclerotin"

/codon_start=1

/protein_id="BAB30678.1"

/db_xref="GI:12856465"

/db_xref="MGI:1921749"

/translation="MPSLAPLICILVHAACVAVGQMOAFRNDATVIFGLGEP

EPPENNQNTNRKENGSRPPHHYDKVSEVSCREIHYTRLTTPGASAPRVELY

CSQCCQPARLLPAIRKVRKWRPNQDFDIPRTRAQVQLCPGASAPRKRRLV

ASCKCRKLTFTFHSGLKDFGPETAPQGRKRPGRAGKANQAELENAY"

BASE COUNT 417 a 565 c 589 g 419 t

ORIGIN

Query Match 100.0%; Score 636; DB 11; Length 1990;

Best Local Similarity 100.0%; Pred. No. 4.4e-127; Indels 0; Gaps 0;

Matches 636; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGCAGCCCTCAGTACGAGCCCGTGCCTCATCTGCTTGGCAGCGCTTCTGTCT 60

44 ATGCAGCCCTCAGTACGAGCCCGTGCCTCATCTGCTTGGCAGCGCTTCTGTCT 103

61 GTGAGGCGCCAGGGGTGGAGCAGCTTCAAGATATGTCACAGAGTCAATCCAGGCTT 120

104 GTGAGGCGCCAGGGGTGGAGCAGCTTCAAGATATGTCACAGAGTCAATCCAGGCTT 163

121 GTGAGGCGCCAGGGGTGGAGCAGCTTCAAGATATGTCACAGAGTCAATCCAGGCTT 180

164 GTGAGGCGCCAGGGGTGGAGCAGCTTCAAGATATGTCACAGAGTCAATCCAGGCTT 223

181 GTGAGGCGCCAGGGGTGGAGCAGCTTCAAGATATGTCACAGAGTCAATCCAGGCTT 240

224 GTGAGGCGCCAGGGGTGGAGCAGCTTCAAGATATGTCACAGAGTCAATCCAGGCTT 283

241 GTGAGGCGCCAGGGGTGGAGCAGCTTCAAGATATGTCACAGAGTCAATCCAGGCTT 300

284 GTGAGGCGCCAGGGGTGGAGCAGCTTCAAGATATGTCACAGAGTCAATCCAGGCTT 343

301 GTGAGGCGCCAGGGGTGGAGCAGCTTCAAGATATGTCACAGAGTCAATCCAGGCTT 360

344 GTGAGGCGCCAGGGGTGGAGCAGCTTCAAGATATGTCACAGAGTCAATCCAGGCTT 403

361 GTGAGGCGCCAGGGGTGGAGCAGCTTCAAGATATGTCACAGAGTCAATCCAGGCTT 420

404 GTGAGGCGCCAGGGGTGGAGCAGCTTCAAGATATGTCACAGAGTCAATCCAGGCTT 463

421 GTGAGGCGCCAGGGGTGGAGCAGCTTCAAGATATGTCACAGAGTCAATCCAGGCTT 480

464 GTGAGGCGCCAGGGGTGGAGCAGCTTCAAGATATGTCACAGAGTCAATCCAGGCTT 523

481 GTGAGGCGCCAGGGGTGGAGCAGCTTCAAGATATGTCACAGAGTCAATCCAGGCTT 540

524 GTGAGGCGCCAGGGGTGGAGCAGCTTCAAGATATGTCACAGAGTCAATCCAGGCTT 583

with BamHI and XhoI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda F1C I. Cloning sites, 5' end: SalI, 3' end: BamHI. Host: DH10B.

Location/Qualifiers

1. 1990

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="PANTOM DB:5430411E23"

/db_xref="MGI:1907679"

/db_xref="taxon:10090"

/clone="5430411E23"

/issue_type="head"

/clone_lib="RIKEN full-length enriched mouse cDNA library"

/dev_stage="6 days neonate"

1. 1990

/gene="Soat"

/gene="Soat"

/note="data source:MGI, source key:MGI.1921749, evidence:ISS"

putative

sclerotin"

/codon_start=1

/protein_id="BAB30678.1"

/db_xref="GI:12856465"

/db_xref="MGI:1921749"

/translation="MPSLAPLICILVHAACVAVGQMOAFRNDATVIFGLGEP

EPPENNQNTNRKENGSRPPHHYDKVSEVSCREIHYTRLTTPGASAPRVELY

CSQCCQPARLLPAIRKVRKWRPNQDFDIPRTRAQVQLCPGASAPRKRRLV

ASCKCRKLTFTFHSGLKDFGPETAPQGRKRPGRAGKANQAELENAY"

Oy	541	GGGCGGAGACCGCGCGGCCGCGAGAAAGGTGCGCAAGCGCGGCCCGCGGCGCGGAGACCC	600	
Db	584	GGCGCGGAGACCGCGCGGCCCGCGAGAAAGGTGCGCAAGCGCGGCCCGCGGCGCGGAGACCC	643	
Oy	601	AAAGCCAAACGAGCGGAGCTGGAGAACCGCTACTAG	636	
Db	644	AAAGCCAAACGAGCGGAGCTGGAGAACCGCTACTAG	679	
RESULT 2				
LOCUS	B636457			
DEFINITION	B636457 RIKEN full-length enriched, adult male aorta and vein Mus musculus cDNA A530001M15 5', mRNA sequence.	667 bp	linear EST 26-Oct-2001	
ACCESSION	B636457			
VERSION	B636457.1	GI:16472326		
KEYWORDS	EST.			
SOURCE	house mouse.			
ORGANISM	Mus musculus			
REFERENCE	Ekakizato; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
AUTHORS	Arakawa,T., Carinci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,U., Komoto,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tsgami,T., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M. and Hayashizaki,Y.			
TITLE	RIKEN Mouse ESTs (Arakawa,T., et al. 2001)			
JOURNAL	Unpublished (2001)			
COMMENT	Contact: Yoshihide Hayashizaki Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute The Institute of Physical and Chemical Research (RIKEN) 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan Tel: 81-45-503-9222 Fax: 81-45-503-9216 Email: genome-resesc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/ Carinci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Komoto,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000) wagii,K., Fujiwaka,S., Inoue,K., Togawa,Y., Iwawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,U., Okazaki,Y., Muramatsu,M., Inoue,Y., Kita,A. and Hayashizaki,Y. RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000) Komoto,H., Fukunishi,Y., Shibata,K., Itoh,M., Carinci,P., Sugahara,Y. and Hayashizaki,Y. Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001) Kondo,S., Shinagawa,A., Saito,T., Kiyasawa,H., Yamanae,I., Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawasoji,Y., Shibata,K. and Hayashizaki,Y. Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001) Please visit our web site (http://genome.gsc.riken.go.jp) for further details. e mouse tissues.			
FEATURES				
SOURCE	location/Qualifiers			
	1..667			
	/organism="Mus musculus"			
	/db_xref="taxon:10090"			
	/clone="A530001M15"			
	/clone_id="RIKEN full-length enriched, adult male aorta and vein"			
	/sex="male"			

```

/tissue_type="aorta and vein"
/dev_stage="adult"
/lab_host="DH10B"
/node="Site_1; Salt; Site_2; BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGGAGAAGATCCAAAGACCTCTTTTTTTTNNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 20.0 and subtraction to Rot = 459.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGGAGAAGATTCGTGAATTATAATTAATCCCCCCCCC 3']. cDNA was cleaved with XhoI and BamHI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda
```

BASE COUNT ORIGIN	119 a	234 c	212 g	100 t	2 others
Query Match	95.6%	Score 608;	DB 10;	Length 667;	
Best Local Similarity	98.1%;	Pred. No. 4.1e-121;	Mismatches 0;	Indels 0;	Gaps 0;
Matches 614;	Conservative				
OY	1 ATGCAGGCCCTCACTAAGCCCGGGCCCTCATCTGCGCTACTTGTCAGCGCTGCGCTTCTGTCT	60			
Dd	41 ATGCAGGCCCTCACTAAGCCCGGGCCCTCATCTGCGCTACTTGTCAGCGCTGCGCTTCTGTCT	100			
OY	61 GTGAGGGCCAGGGGTGCGAACCTTTCAGAAATGATVGCACAGAGTCATCCAGGGCTT	120			
Dd	101 GTGAGGGCCAGGGGTGCGAACCTTTCAGAAATGATVGCACAAAGTCATCCAGGGCTT	160			
OY	121 GAGAGTACCCCAGACCCTCTCTGAGAAACAACAGACATGAACCGGGCCGAGAAATGA	180			
Dd	161 GAGAGTACCCCAGACCCTCTCTGAGAAACAACAGACATGAACCGGGCCGAGAAATGA	220			
OY	181 GGACAGACCTCCCAACATCCTTAGACCGCAAAGATGTGTCCAGTAGAGTGC CGCGAG	240			
Dd	221 GGACAGACCTCCCAACATCCTTAGAGCCCAAGATGTGTAGAGTAGACGTGC CGCGAG	280			
OY	241 CTGCACTACACCCGTTTCTTGA CAGACGGCCCATGCGCAGAGGCCAAGCCGGTCA CCGAG	300			
Dd	281 CTGCACTACACCCGTTTCTTGA CAGACGGCCCATGCGCAGAGGCCAAMCCGGTCA CCGAG	340			
OY	301 TTGGTGTCTCCGGCAGTGTGCGGCCCGCGCGCTGCTGCCCAACGCATCGGGCGCGT	360			
Dd	341 TTGGTGTCTCCGGCAGTGTGCGGCCCGCGCGCTGCTGCCCAACGCATCGGGCGCGT	400			
OY	361 AAGTGTGGCGCCGAAACGAGACCGGATTTCCGTCGATCCCGGATCGCTACCGCGGAG	420			
Dd	401 AAGTGTGGCGCCGAAACGAGACCGGATTTCCGTCGATCCCGGATCGCTACCGCGGAG	460			
OY	421 CGGGTCAGCTGTGTGCCCCGGGGGGCGCGCGCGCTGGCGAAAGTGCGTGTGGTG	480			
Dd	461 CGGGTCAGCTGTGTGCCCCGGGGGGCGCGNGCGCGCTGGCGAAAGTGCGTGTGGTG	520			
OY	481 GCCTGTCAAGTGCAACGCGCTCACCCGCTTTCACAACAACAGTCGAGCTCAAGACTTC	540			
Dd	521 GCCTGTCAAGTGCAAGTGCACCGCTTTCACAACAACAGTCGAGCTTCAAGACTTC	580			
OY	541 GGGCCGAGACCGCGCGCGCGCGAGAGGCTGCAAGCTCGCGACCCCGCGCGGAGACC	600			
Dd	581 GGGCCGAGAACCGCGCGCGCGCGCAAAATGTGTGCAAGCTCGCGACCCCGCGGAGACC	640			
OY	601 AAAGCCAACAGCGGCGAGCTGAGAA	626			
Dd	641 AAAGCCAACAGCGGCGAGCTTGAAGAA	666			

RESULT 3

	B637315	623 bp	mRNA	linear	EST 26-OCT-2001	
DEFINITION	B637315 RIKEN full-length enriched, adult male aorta and vein Mus musculus cDNA clone A530065C22 5', mRNA sequence.					
ACCESSION	B637315					
VERSIONS	B637315.1 GI:16473100					
SOURCE	house mouse. Mus musculus					
ORANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.					
REFERENCE	1 (bases 1 to 623)					
AUTHORS	Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Himoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Komno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M. and Hayashizaki.Y.					
TITLE	RIKEN Mouse ESTs (Arakawa,T., et al. 2001)					
JOURNAL	Unpublished (2001)					
COMMENT	Contact: Yoshihide Hayashizaki Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute The Institute of Physical and Chemical Research (RIKEN) 1-7-22 Suito-cho, Tsunumi-ku, Yokohama, Kanagawa 230-0045, Japan Tel.: 81-45-503-9222 Fax: 81-45-503-9216 Email: genome-res@sc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/ Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Komno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000) wegi,K., Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watabiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000) Komno,H., Fukuishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,Y. and Hayashizaki,Y. Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001) Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanka,I., Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and Hayashizaki,Y. Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001) Please visit our web site (http://genome.gsc.riken.go.jp) for further details. e mouse tissues.					
FEATURES	Location/Qualifiers					
SOURCE	1..623 /organism="Mus musculus" /db_xref="taxon:10090" /clone="A530065C22" /clone_jib="RIKEN full-length enriched, adult male aorta and vein" /sex="male" /tissue_type="aorta and vein" /dev_stage="adult" /lab_host="DH10B" /note="Site_1: SalI; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5'					

	GAGGAGAGAAGATCCAGAGCTTTTTTTTTTTTTVN 3'), cDNA was prepared by using trehlose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 20.0 and subtraction to Rot = 459.0. Second strand cDNA was prepared with the primer adapter of sequence 15' GAGGAGAGATTCTCGAATTAATTAATACCCCCCCCCC 3'. cDNA was cleaved with XhoI and BamHI. Vector: a modified pBluescript KS(+) after bulk excision from lambda
BASE COUNT	106 a 222 c 199 g 96 t
ORIGIN	FLC I."
Query Match	91.7%; Score 583; DB 10; Length 623;
Best Local Similarity	100.0%; Pred. No. 9.7e-116;
Matches 583; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
OY	1 ATGCAGCCCTCACTAGAGCCCGTGCTCATCTGCCTACTTGTGACGGCTGCCCTTCGTGCT 60
DB	41 ATGCAGCCCTCACTAGAGCCCGTGCTCATCTGCCTACTTGTGACGGCTGCCCTTCGTGCT 100
OY	61 GTGAGGGCCAGAGGGGTGGCANAGCTTCTAGAAATGATGCCACAGAGTATCCAGGGCTT 120
DB	101 GTGAGGGCCAGAGGGGTGGCAAAGCTTCTAGAAATGATGCCACAGAGTATCCAGGGCTT 160
OY	121 GGAAGTACCCCAGAGCCCTCTCTGAGAACCAACAAGAACCCGGGGGAGAGATGGA 180
DB	161 GGAAGTACCCCAGAGCCCTCTCTGAGAACCAACAAGAACCCGGGGGAGAGATGGA 220
OY	181 GGACAGACTCCCAACCATCTTATGACGCCAAGATGTGTCCAGTACAGCTGCGCGAG 240
DB	221 GGACAGACTCCCAACCATCTTATGACGCCAAGATGTGTCCAGTACAGCTGCGCGAG 280
OY	241 CTGCATCACACCCTCGCTTCTTAGACAGACGGCCATGCGCAGCGCCACCGGTACCGAG 300
DB	281 CTGCATCACACCCTCGCTTCTTAGACAGACGGCCATGCGCAGCGCCACCGGTACCGAG 340
OY	301 TTGGTGTGCTCCGGTCAATGAGGGGCCCGGGGCTGTGCTGCCAAGCCATGGGCGCGTG 360
DB	341 TTGGTGTGCTCCGGTCAATGAGGGGCCCGGGGCTGTGCTGCCAAGCCATGGGCGCGTG 400
OY	361 AAGTGTGGCGCCCGAAGCGAACCGGATTTCCGCTGCATCCCGATCGCTACCGCGCGAG 420
DB	401 AAGTGTGGCGCCCGAAGCGAACCGGATTTCCGCTGCATCCCGATCGCTACCGCGCGAG 460
OY	421 CGGGTGCAGCTGCTGTGCCCCGGGGGCGCGCGCGCTCGCGCAAGGTGCTGTGTTG 480
DB	461 CGGGTGCAGCTGCTGTGCCCCGGGGGCGCGCGCGCTCGCGCAAGGTGCTGTGTTG 520
OY	481 GCCTGTGTCAAGTGCNAAGCGCTCACCCGCTTCCACACCAAGTGGAGCTCAAGACTTC 540
DB	521 GCCTGTGTCAAGTGCNAAGCGCTCACCCGCTTCCACACCAAGTGGAGCTCAAGACTTC 580
OY	541 GGGCGCGAGACCGCGCGGCGCGCAGAGAAGGTGCGCAAGCGCGGAGC 583
DB	581 GGGCGCGAGACCGCGCGGCGCGCAGAGAAGGTGCGCAAGCGCGGAGC 623
RESULT 4	
BB638050	511 bp mRNA linear EST 31-AUG-2001
LOCUS	
DEFINITION	BB638050 RIKEN full-length enriched, 3 days neonate thymus Mus
ACCESSION	musculus cDNA clone AB30006E24 5', mRNA sequence.
VERSION	BB638050
KEYWORDS	BB638050.1 GI:15401086
SOURCE	EST.
ORGANISM	house mouse.
MUS	mus musculus
Eukaryota;	Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia;	Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
(bases 1 to 511)	
Arakawa,T.,	Carinini,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.,
Hiramoto,K.,	Hori,F., Ishii,Y., Ito,M., Kawai,U., Komio,H., Koude,

TITLE
JOURNAL
COMMENT

M., Koya, S., Matsumura, T., Miyazaki, A., Nomura, K., Ohno, M., Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toya, T., Muramatsu, M., and Hayashizaki, Y.
RIKEN Mouse ESTs (Arai, T., et al. 2001)
Unpublished (2001)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@sc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. *Genome Res.* 10 (10), 1617-1630 (2000)
Wagl, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, B., Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A., and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multichannel sequencer. *Genome Res.* 10 (11), 1757-1771 (2000)
Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara, Y., and Hayashizaki, Y.
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. *Genome Res.* 11 (2), 281-289 (2001)
Yamanaka, I., Kiyosawa, H., Kondo, S., Saito, T., Shinagawa, A., Aizawa, K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K., Arai, T., Ishii, Y., and Hayashizaki, Y.
Mapping of 19032 mouse cDNAs on mouse chromosomes. *J. Struct. Funct. Genomics* 2 pre, L72-L86 (2001)
Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.
e mouse tissues.

FEATURES
source

Location/Qualifiers
1. 511
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="A63006E24"
/clone_1ib="RIKEN full-length enriched, 3 days neonate thymus"
/tissue_type="thymus"
/dev_stage="3 days neonate"
/lab_host="DH10B"
/note="Site 1: Salt; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5'
GAGAGAGAGAGATCCAGACCTCTTTTCTTTTCTTTT 3']. cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 20.0 and subtraction to Rot = 459.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGATCTCCAGCTTAAATTAATCCCCCCCC 3']. cDNA was cleaved with XhoI and BamHI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I."

BASE COUNT
ORIGIN

86 a 182 c 162 g 81 t

Query Match 73.8%; Score 469.4; DB 10; Length 511;
Best Local Similarity 99.8%; Pred. No. 2.7e-91;

Matches 470; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGCAGCCCTCACTAGCCCGGCTCTATCTGCTTGTGACGCTGCTGCT 60
Db 41 ATGCAGCCCTCACTAGCCCGGCTCTATCTGCTTGTGACGCTGCTGCT 100
Qy 61 GTGAGGCGCAGAGGGGTGCGAAGCTTTCAGAAATGTCACAGAGTCTCCAGGGCTT 120
Db 101 GTGAGGCGCAGAGGGGTGCGAAGCTTTCAGAAATGTCACAGAGTCTCCAGGGCTT 160
Qy 121 GGAGAGTACCCCGAGAGCTCTCTGGAAGAACACAGACCAATGAACCGGGCGAAGTGA 180
Db 161 GGAGAGTACCCCGAGAGCTCTCTGGAAGAACACAGACCAATGAACCGGGCGAAGTGA 220
Qy 181 GGCAAGCTCCCGACCATCTCTATGACCGCAAGATGTCTCCAGTACAGCTGCGCGAG 240
Db 221 GGCAAGCTCCCGACCATCTCTATGACCGCAAGATGTCTCCAGTACAGCTGCGCGAG 280
Qy 241 CTGCACTACACCCGCTTCTTGAACAGCGCCCATGCGCAGCGCCGATCACCAG 300
Db 281 CTGCACTACACCCGCTTCTTGAACAGCGCCCATGCGCAGCGCCGATCACCAG 340
Qy 301 TTGGTGTGCTCGGCGCATGCGCGCCCGCGCTGCTGCTCCCAAGCCATGCGCGCTG 360
Db 341 TTGGTGTGCTCGGCGCATGCGCGCCCGCGCTGCTGCTCCCAAGCCATGCGCGCTG 400
Qy 361 AAGTGTGCGCGCGCGAAGCGAAGCGAATTTCCGCTGATCCCGATCGCTACCGCGCG 420
Db 401 AAGTGTGCGCGCGCGAAGCGAAGCGAATTTCCGCTGATCCCGATCGCTACCGCGCG 460
Qy 421 CGGTCGAGCTGCTGTGCGCGCGCGCGCGCGCGCTGCGCGAAGTGTG 471
Db 461 CGGTCGAGCTGCTGTGCGCGCGCGCGCGCGCGCGCTGCGCGAAGTGTG 511

RESULT 5

BE101082/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

EST

Rattus

REFERENCE

AUTHORS

TITLES

JOURNAL

MEDLINE

COMMENT

BE101082 419 bp mRNA linear EST 13-JUN-2000
UI-R-BJ1-aly-h-10-0-UI.s1 UI-R-BJ1 Rattus norvegicus cDNA clone
UI-R-BJ1-aly-h-10-0-UI 3', mRNA sequence.
BE101082
BE101082.1 GI:8492989
EST.
Norway rat.
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 419)
Bonaldo, M.F., Lennon, G., and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
97044477
Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: msoares@iowa.uiowa.edu
Oligo-dT track not found. Not 1 site shown in beginning of sequence
is likely internal to the message. cDNA Library Preparation: M.B.
Soares lab clone distribution: clones will be available through
Research Genetics (www.resgen.com) The following repetitive
elements were found in this cDNA sequence: 97-125,
>GC-richlow_complexity
Seg primer: M13 Forward
POLYA=NO.

FEATURES
source

Location/Qualifiers

1. 419
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"

/db_xref="taxon:10116"
 /clone="UI-R-BJ1-acy-h-10-0-UI"
 /clone_1lb="UI-R-BJ1"
 /lab_host="DH10B (Life Technologies)"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-BJ1
 library is a subtracted library derived from the following
 tissues: atrium at 16.5 dpc, ventricle at 16.5 dpc, AV
 canal at 16.5 dpc, atrium at 15 dpc, ventricle at 15 dpc,
 AV canal at 15 dpc, ventricle at 13 dpc, and adult heart.
 For a detailed description of the library from which this
 clone was derived, please visit our web site at
 ratest.eng.uiowa.edu. The subtraction has been previously
 described in (Bonaldo, Lennon and Soares, Genome Research
 6:791-806, 1996)
 Tag_SEQ=None found"
 BASE COUNT 56 a 135 c 155 g 73 t
 ORIGIN

Query Match 61.5%; Score 391.4; DB 10; Length 419;
 Best Local Similarity 96.2%; Pred. No. 1.7e-74;
 Matches 401; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 145 GAGAACACAGACATGAACCGGGCGAGATGAGGACGACCTCCCACTAT 204
 DB 417 GAGAACACAGACATGAACCGGGCGAGATGAGGACGACCTCCCACTAT 358
 QY 205 GAGCCCAAGATGTGTCCGATGACGTCCCGGAGCTGCACTACCCGCTTCTGACA 264
 DB 357 GACACCAAGATGTGTCCGATGACGTCCCGGAGCTGCACTACCCGCTTCTGACC 298
 QY 265 GAGGCCCATCGGAGCGGACCGGATGACCGATTTGGTGTCTCGGCGAGTGGCG 324
 DB 297 GAGGCCCATCGGAGCGGACCGGATGACCGATTTGGTGTCTCGGCGAGTGGCG 238
 QY 325 CCCGCGGCTGTGCTCCCAACGCGGATGAGGATGAGTGGCGCGCCGGAACGACG 384
 DB 237 CCCGCGGCTGTGCTCCCAACGCGGATGAGGATGAGTGGCGCGCCGGAACGACG 178
 QY 385 GATTTCGCTGATCCGAGATGCTACCGCGGAGCGGGTGTGCTGTGCTGCTGCG 444
 DB 177 GACTTCGCTGATCCGAGATGCTACCGCGGAGCGGGTGTGCTGTGCTGCTGCG 118
 QY 445 GGGCGGGCGGCGGCTGCGGCAAGTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 504
 DB 117 GGGCGGGCGGCGGCTGCGGCAAGTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 58
 QY 505 ACCGCTTCCCAACCACTGAGCTCAAGACTTTCGAGCGGAGACCGGCGCG 561
 DB 57 ACCGCTTCCCAACCACTGAGCTCAAGACTTTCGAGCGGAGACCGGCGCG 1

RESULT 6
 A1556282/c 360 bp mRNA linear EST 23-MAR-1999
 LOCUS
 DEFINITION UI-R-C2p-rk-a-03-0-UI s1 UI-R-C2p Rattus norvegicus cDNA clone
 ACCSSION UI-R-C2p-rk-a-03-0-UI 3', mRNA sequence.
 VERSION A1556282.1 GI:4488645
 KEYWORDS EST.
 SOURCE Norway rat.
 ORGANISM Rattus norvegicus
 Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE 1 (bases 1 to 360)
 AUTHORS Bonaldo, M.F., Lennon, G. and Soares, M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene
 JOURNAL discovery
 MEDLINE Genome Res. 6 (9), 791-806 (1996)
 COMMENT
 Contact: Soares, MB
 Program for Rat Gene Discovery and Mapping

University of Iowa
 451 Eckstein Medical Research Building Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: msoares@blue.weeg.uiowa.edu
 Oligo-dT track not found. Not I site shown in beginning of sequence
 is likely internal to the message. cDNA library preparation: M.
 Fatima Bonaldo, Ph.D. Clone distribution: clones will be available
 through Research Genetics (www.resgen.com) The following repetitive
 elements were found in this cDNA sequence: 97-125,
 >GC-rich#low-complexity
 Seq primer: M13 Forward.
 Location/Qualifiers
 1..360
 /organism="Rattus norvegicus"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /clone="UI-R-C2p-rk-a-03-0-UI"
 /clone_1lb="UI-R-C2p"
 /dev_stage="adult"
 /lab_host="DH10B (Life Technologies)"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-C2p
 library is a subtracted library derived from the UI-R-C1
 library, which is a subtracted library derived from the
 UI-R-C0 library. The UI-R-C0 library consisted of a
 mixture of individually tagged normalized libraries
 constructed from rat placenta, adult lung, brain, liver,
 kidney, heart, spleen, ovary, muscle, 8, 12 and 18-day
 embryo. The tag is a string of 3-5 nucleotides present
 between the Not I site and the oligo-dT track which allows
 identification of the library of origin of a clone within
 the mixture. The subtracted library (UI-R-C2p) was
 constructed as follows: PCR amplified cDNA inserts from
 UI-R-C1 clones from which 3' ESTs had been derived was
 used as a driver in a hybridization with the UI-R-C1
 library in the form of single-stranded circles. The
 remaining single-stranded circles (subtracted library) was
 purified by hydroxyapatite column chromatography,
 converted to double-stranded circles and electroporated
 into DH10B bacteria (Life Technologies) to generate the
 UI-R-C2p library. This procedure has been previously
 described (Bonaldo, Lennon and Soares, Genome Research 6:
 791-806, 1996)"

BASE COUNT 54 a 120 c 131 g 54 t 1 others
 ORIGIN

Query Match 52.9%; Score 336.6; DB 9; Length 360;
 Best Local Similarity 95.8%; Pred. No. 1e-62;
 Matches 345; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 202 TATGACGCCAAGATGTGTCCGATGACGCTCCGCGAGCTGCACTACCCGCTTCTG 261
 DB 360 TATGACGCCAAGATGTGTCCGATGACGCTCCGCGAGCTGCACTACCCGCTTCTG 301
 QY 262 ACAGACGGCCCAATGCGGAGCGGCAAGCGGATGAGTGTGCTGCGGACAGTGC 321
 DB 300 ACAGACGGCCCAATGCGGAGCGGCAAGCGGATGAGTGTGCTGCGGACAGTGC 241
 QY 322 GGGCGGCGGCGGCTGCGGCAAGCGGATGAGTGTGCTGCGGACAGTGC 381
 DB 240 GGGCGGCGGCGGCTGCGGCAAGCGGATGAGTGTGCTGCGGACAGTGC 181
 QY 382 CGGATTTCCGCTGATCCCGGATGCTACCGCGGAGCGGGTGTGCTGTGCTGCTG 441
 DB 180 CGGATTTCCGCTGATCCCGGATGCTACCGCGGAGCGGGTGTGCTGTGCTGCTG 121
 QY 442 GGGGGGCGGCGGCGGCTGCGGCAAGTGTGCTGAGTGTGCTGCTGCTGCTGCTG 501
 DB 120 GGGGGGCGGCGGCGGCTGCGGCAAGTGTGCTGAGTGTGCTGCTGCTGCTGCTG 61
 QY 502 CTCACCGCTTCCCAACCACTGAGCTCAAGACTTTCGAGCGGAGACCGGCGCG 561

CC and designated as Cloaked-2, derived from human and mouse. The cloaked-2
 CC polypeptides can be expressed by standard recombinant methodology. The
 CC cloaked-2 polynucleotides are useful in gene therapy and antisense
 CC therapy. The cloaked-2 polypeptides and polynucleotides are useful for
 CC treating, preventing, ameliorating or detecting diseases and disorders of
 CC the kidney (e.g. anemia, hypertension or low blood pressure), heart (e.g.
 CC cardiac hypertrophy, congestive heart failure, myocardial infarction,
 CC arrhythmias, atherosclerosis, hypertension or low blood pressure),
 CC skeletal muscle (e.g. muscular dystrophy or cachexia), placenta (e.g.
 CC congenital abnormalities or miscarriage), liver (e.g. hepatitis or
 CC cirrhosis), pancreas (e.g. diabetes or pancreatitis), thyroid (e.g.
 CC Grave's disease or myxedema) or adrenal cortex (e.g. Cushing's disease
 CC or Addison's disease), homeostasis or metabolic diseases (e.g. obesity,
 CC cancer or myopathies), infections, or autoimmune diseases. Selective
 CC binding agents may be used to modulate the biological activities of
 CC cloaked-2 polypeptides or to detect cloaked-2 polypeptide levels in a
 CC sample. Transgenic non-human animals are useful for drug candidate
 CC screening. The present sequence represents the mouse cloaked-2 mature
 CC polypeptide.

CC Sequence 188 AA;

Query Match 90.7%; Score 951; DB 23; Length 188;

Best Local Similarity 89.5%; Pred. No. 4.6e-84; Mismatches 8; Indels 2; Gaps 1;

Matches 170; Conservative 10; Mismatches 8; Indels 2; Gaps 1;

1 OGMQAFKNDATETIIPELGEYEPPELENNKTNRAENGSRPHHPFTKVSEYSCREL 60
 1 OGMQAFKNDATETIIPELGEYEPPELENNKTNRAENGSRPHHPFTKVSEYSCREL 58

61 HFTRYVTGPGCSAKPVTELVCSGGCPARLLPNAIGRKWMPSPGPPFCIPDRYRQR 120
 59 HFTRYVTGPGCSAKPVTELVCSGGCPARLLPNAIGRKWMPSPGPPFCIPDRYRQR 118

121 VQLLCGGEAPRRKRVLVASCKCKRLTRFNOSLKDFTGEARPOKGRKPRPARSAK 180
 119 VQLLCGGEAPRRKRVLVASCKCKRLTRFNOSLKDFTGEARPOKGRKPRPARSAK 178

181 ANQAELENAY 190
 179 ANQAELENAY 188

Db 179 ANQAELENAY 188

RESULT 14
 ID AAY96432
 AAY96432 standard; Protein; 211 AA.

AC AAY96432;

DT 12-SEP-2000 (first entry)

DE Murine TGF-beta binding protein (BBER).

KM osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
 KM BBER; gene therapy; antisense therapy; fracture; bone mineralization.

OS Mus musculus.

PN WO200032773-A1.

PD 08-JUN-2000.

PF 24-NOV-1999; 99WO-US27990.

PR 27-NOV-1998; 98US-0110283.

PA (DARN-) DARWIN DISCOVERY LTD.

PI Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepfer BW;

PI Van Ness J, Winkler DG;

DR WPI, 2000-412321/35.

DR N-PSDB; AAA29058.

XX Nucleic acids (1) encoding a transforming growth factor beta binding
 PT protein, useful for identifying agents for treating osteopenia,
 PT osteoporosis and fractures

PS Claim 5; Page 124; 162pp; English.

XX This shows a murine transforming growth factor-beta (TGF-beta)
 CC binding protein designated mBER. The cDNA and protein may be used for
 CC prevention, treatment and diagnosis of diseases associated with
 CC inappropriate BBER expression. For example, they may be used to treat
 CC disorders associated with decreased TGF-beta BP expression. The cDNA or
 CC vectors may be administered to treat diseases by rectifying mutations or
 CC deletions in a patient's genome that affect the activity of BBER by
 CC expressing inactive proteins or to supplement the patients own production
 CC of BBER polypeptides. The nucleic acids may be used for recombinant
 CC production of BBER, gene therapy, antisense therapy, as probes for
 CC diagnostic assays and for functional studies. BBER may be used to raise
 CC antibodies and for identification of BBER modulators. BBER antagonists
 CC may be used to increase bone mineral content for the treatment of
 CC disorders such as osteopenia, osteoporosis, fractures and other disorders
 CC associated with low mineral content.

XX Sequence 211 AA;

Query Match 90.7%; Score 951; DB 21; Length 211;

Best Local Similarity 89.5%; Pred. No. 5.2e-84; Mismatches 8; Indels 2; Gaps 1;

Matches 170; Conservative 10; Mismatches 8; Indels 2; Gaps 1;

1 OGMQAFKNDATETIIPELGEYEPPELENNKTNRAENGSRPHHPFTKVSEYSCREL 60
 24 OGMQAFKNDATETIIPELGEYEPPELENNKTNRAENGSRPHHPFTKVSEYSCREL 81

61 HFTRYVTGPGCSAKPVTELVCSGGCPARLLPNAIGRKWMPSPGPPFCIPDRYRQR 120
 82 HFTRYVTGPGCSAKPVTELVCSGGCPARLLPNAIGRKWMPSPGPPFCIPDRYRQR 141

121 VQLLCGGEAPRRKRVLVASCKCKRLTRFNOSLKDFTGEARPOKGRKPRPARSAK 180
 142 VQLLCGGEAPRRKRVLVASCKCKRLTRFNOSLKDFTGEARPOKGRKPRPARSAK 201

181 ANQAELENAY 190
 202 ANQAELENAY 211

Db 202 ANQAELENAY 211

RESULT 15
 ID ABB07210
 ABB07210 standard; Protein; 211 AA.

AC ABB07210;

DT 26-MAR-2002 (first entry)

DE Mouse cloaked-2 polypeptide sequence.

KM Cloaked-2; cysteine knot motif; nephrotropic; cardiac; immunomodulator;
 KM hepatotropic; antiinflammatory; antihypertoid; cytoprotective;
 KM antianemic; hypotensive; antiarrhythmic; antidiabetic; muscular;
 KM antidiabetic; anorectic; gene therapy; cell therapy; antisense therapy;
 KM mouse.

OS Mus musculus.

PH Key

FT Peptide

FT Protein

FT Misc-difference 198..200 /note= "mature protein (ABB07208) "

FT /note= "this region is missing in the sequence provided
 in the sequence listing but has been indicated
 correctly in the sequence in the Figure"

```
XX MO200192308-A2.
PN
XX
XX 06-DEC-2001.
PD
XX 29-MAY-2001; 2001WO-US17478.
PF
XX 01-JUN-2000; 2000US-208550P.
PR
XX 04-AUG-2000; 2000US-223542P.
XX
PA (AMGE-) AMGEN INC.
XX
XX Paszty CJ, Gao Y;
PI
XX
XX WPI; 2002-114325/15.
XX
XX DR N-PSDB; ABA94294.
XX
XX New human and mouse cysteine-knot polypeptide designated as Cloaked-2,
PT for treating or preventing kidney, heart (e.g. myocardial infarction)
PT or liver (e.g. hepatitis) diseases
XX
XX
XX Example 2; Fig 2; 170pp; English.
PS
XX
XX The invention relates to polypeptides comprising a cysteine knot motif
CC and designated as Cloaked-2, derived from human and mouse. The Cloaked-2
CC polypeptides can be expressed by standard recombinant methodology. The
CC Cloaked-2 polynucleotides are useful in gene therapy and antisense
CC therapy. The Cloaked-2 polypeptides and polynucleotides are useful for
CC treating, preventing, ameliorating or detecting diseases and disorders of
CC the kidney (e.g. anemia, hypertension or low blood pressure), heart (e.g.
CC cardiac hypertrophy, congestive heart failure, myocardial infarction,
CC arrhythmias, atherosclerosis, hypertension or low blood pressure),
CC skeletal muscle (e.g. muscular dystrophy or cachexia), placenta (e.g.
CC congenital abnormalities or miscarriage), liver (e.g. hepatitis or
CC cirrhosis), pancreas (e.g. diabetes or pancreatitis), thyroid (e.g.
CC Grave's disease or myxedema) or adrenal cortex (e.g. Cushing's disease
CC or Addison's disease), homeostasis or metabolic diseases (e.g. obesity,
CC cancer or myopathies), infections, or autoimmune diseases. Selective
CC binding agents may be used to modulate the biological activities of
CC Cloaked-2 polypeptides or to detect Cloaked-2 polypeptide levels in a
CC sample. Transgenic non-human animals are useful for drug candidate
CC screening. The present sequence represents the mouse Cloaked-2
CC polypeptide.
XX
XX
SQ Sequence 211 AA;
Query Match 90.7%; Score 951; DB 23; Length 211;
Best Local Similarity 89.5%; Pred. No. 5.2e-84;
Matches 170; Conservative 10; Mismatches 8; Indels 2; Gaps 1;
QY 1 QGWAQAFKNDATETIPELGEYEPPELLENNKTKRAENGRRPHHPETKDVSEYSCREL 60
Db 24 QGWAQAFKNDATETIPELGEYEPPELLENNKTKRAENGRRPHHPETKDVSEYSCREL 81
QY 61 HFTRYVTDDGPCRSKAPVTELVCSGCCGPARLLPNAIGRGKWRBSPDPFRICIPDRYRAQR 120
Db 82 HFTRYVTDDGPCRSKAPVTELVCSGCCGPARLLPNAIGRGKWRBSPDPFRICIPDRYRAQR 141
QY 121 VOLLCPGGEARPRARKRLVASCCKRLTRFNQSELKDPGTEARPOKGRKPRPRASAK 180
Db 142 VOLLCPGGEARPRARKRLVASCCKRLTRFNQSELKDPGTEARPOKGRKPRPRASAK 201
QY 181 ANOALEENAY 190
Db 202 ANOALEENAY 211
```

Search completed: March 28, 2003, 14:16:13
Job time : 54.2 secs

GenCore version 5.1.4 p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 28, 2003, 14:13:46 ; Search time 27.36 Seconds
(without alignments)
667.600 Million cell updates/sec

Title: US-09-867-274-2

Perfect score: 1049

Sequence: 1 QGMQAFKNDATETIIPELGEV.....KPRPRASAKANQALENAV 190

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	270	25.7	134	2 T08710	hypothetical prote
2	94.5	9.0	159	2 T51373	luteinizing hormon
3	87.5	8.3	158	2 A61091	lutropin beta chai
4	87.5	8.3	601	2 T22025	hypothetical prote
5	87.5	8.3	601	2 D89711	protein P40E10.4 (
6	82.5	7.9	1469	2 B3665	slit protein 2 pre
7	81	7.7	270	2 S71793	head-inducing fact
8	80.5	7.7	715	2 T12534	hypothetical prote
9	80	7.6	1480	2 A3665	slit protein 1 pre
10	79	7.5	118	2 PN0141	lutropin beta chai
11	79	7.5	178	2 T51824	tumor-suppressive
12	79	7.5	178	2 A47291	probable finger pr
13	79	7.5	866	2 T29197	hypothetical prote
14	78.5	7.5	166	2 T51242	luteinizing hormon
15	78.5	7.5	969	2 A70912	probable leus prot
16	78	7.4	1142	2 T30272	hypothetical prote
17	78	7.4	1210	1 GQHUE	epidermal growth f
18	77.5	7.4	1042	2 A57534	mucin 5AC (clone L
19	77	7.3	118	2 PN0139	lutropin beta chai
20	77	7.3	510	2 A42750	insulinoma-associ
21	77	7.3	520	2 S47142	matng type A prot
22	77	7.3	839	2 T50590	class I INCENP pro
23	76.5	7.3	1223	2 S29717	adenylate cyclase
24	76	7.2	398	2 A35281	intergenetary muc
25	76	7.2	2142	2 B35098	MHC class III hist
26	75.5	7.2	403	2 C70832	hypothetical prote
27	75.5	7.2	646	2 T02398	hypothetical prote
28	75.5	7.2	1436	2 A46496	antigen Wc1.1 prec
29	75	7.1	2153	2 T30074	hypothetical prote

30	74.5	7.1	206	2 G87592	hypothetical prote
31	74.5	7.1	824	2 T10020	leucine-tRNA ligas
32	74.5	7.1	972	2 T10023	bifocal protein -
33	74.5	7.1	1063	2 T03743	hypothetical prote
34	74	7.1	269	2 E69381	apolipoprotein H-r
35	74	7.1	343	2 G35070	probable helicase
36	74	7.1	702	2 E72775	class II INCENP pr
37	74	7.1	877	2 T50591	hypothetical prote
38	74	7.1	2090	2 T30075	serum-inducible kl
39	73.5	7.0	682	2 A44493	hypothetical prote
40	73.5	7.0	864	2 H85335	hypothetical prote
41	73.5	7.0	864	2 T04518	hypothetical prote
42	73	7.0	155	2 C95002	U49 protein - hum
43	73	7.0	301	2 WMBEP9	9eg-like protein -
44	73	7.0	613	2 T00077	submaxillary mucin
45	73	7.0	1589	2 T42233	

ALIGNMENTS

RESULT 1

T08710
hypothetical protein DKFZP564D206.1 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 13-Aug-1999
A:Accession: T08710
R:Mandut, R.; Heubner, D.; Mewes, H.W.; Gaassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, March 1999
A:Reference number: Z16471
A:Accession: T08710
A:Molecule type: mRNA
A:Residues: 1-134 <N&M>
A:Cross-references: EMBL:AL050024
A:Experimental source: fetal brain; clone DKFZP564D206
C:Genetics:
A>Note: DKFZP564D206.1

Query Match 25.7% Score 270; DB 2; Length 134;

Best local similarity 44.2%; Pred. NO. 4e-17; Mismatches 57; Conservative 23; Mismatches 43; Indels 6; Gaps 4;

QY	57	CRELHFRVTVDGPRSAKPYTELVCSGCGPARLPNALIGRG--KWM-RPSGDPFRCI	112
DB	3	CRELHFRVTVDGPRSAKPYTELVCSGCGPARLPNALIGRG--KWM-RPSGDPFRCI	62
QY	113	PDRYAGQVOLLCPGEGAPRAKRVRLVASCCKRLTRFHNSLKDRTGARPOKGRKP	172
DB	63	NDKRTORIQLOCDG-STRTYKITVTVACKCKRYTROHNSHNFSPSPAKPVQHRE	121
QY	173	RPRA-RSAK 180	
DB	122	KRAKSSK 130	

RESULT 2

luteinizing hormone beta subunit - turkey
C:Species: Meleagris gallopavo (common turkey)
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 16-Jul-1999
A:Accession: I51373
J:Yon, S.; Foster, L.K.; Silsby, J.L.; el Halawani, M.E.; Foster, D.N.
J Mol. Endocrinol. 14, 117-129, 1995
A>Title: Sequence analysis of the turkey LH beta subunit and its regulation by gonadotropin
A:Reference number: I51373; MUID:95290073; PMID:7772235
A:Accession: I51373
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-159 <YOU>
A:Cross-references: GB:I35519; MID:g530952; PIDN:AAA74125.1; PID:g530953
C:Genetics:
A:Gene: LH-beta
C:Superfamily: pituitary glycoprotein hormone beta chain

Query Match 9.0%; Score 94.5; DB 2; Length 159;
Best Local Similarity 29.4%; Pred. No. 0.25; Indels 21; Gaps 5;
Matches 32; Conservative 7; Mismatches 49; Indels 21; Gaps 5;

QY 39 GGRPHHPEFKDVSEVSCRELFHFRVYTDGPCRSAPVTELVCSGCCGPARLLPNAIGR 98
|||||
DB 43 GGRPP-----CRPINTVAVEKDECPQCMATTTTACGVCRRR-----EPPVYR 85

QY 99 GKMRPSPGDFRCIPDRYRAQVOLL-CPGGEAPRARKVRLVASCCKR 146
DB 86 SPLGRP--POSSCTYGAIRYERMAWMCPIGSDPRV-LLPVALSCRCAR 131

RESULT 3
A61091
Lutropin beta chain precursor - chicken
N/Alternate names: luteinizing hormone beta chain
C/Species: Gallus gallus (chicken)
C/Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 08-Dec-1995
C/Accession: A61091
R/Note: T.; Ando, H.; Ueda, T.; Kubokawa, K.; Higashinakagawa, T.; Ishii, S.
J. Mol. Endocrinol. 3, 129-137, 1989
A/Title: Molecular cloning and nucleotide sequence analysis of the putative cDNA for the
A/Reference number: A61091; MUID:89374710; PMID:2476156
A/Accession: A61091
A/Status: not compared with conceptual translation
A/Molecule type: mRNA
A/Residues: 1-158 <NOC>
C/Superfamily: pituitary glycoprotein hormone beta chain
C/Keywords: glycoprotein; hormone; pituitary
F/1.39/Domains: signal sequence #status predicted <SIG>
F/48-73,62-96,65-127,77-149,111-139,129-132/Disulfide bonds: #status predicted
F/52/Binding site: carbohydrate (asn) (covalent) #status predicted

Query Match 8.3%; Score 87.5; DB 2; Length 158;
Best Local Similarity 27.9%; Pred. No. 1;
Matches 31; Conservative 8; Mismatches 47; Indels 25; Gaps 5;

QY 39 GGRPHHPEFKDVSEVSCRELFHFRVYTDGPCRSAPVTELVCSGCCGPARLLPNAIGR 98
|||||
DB 43 GGRPP-----CRPINTVAVEKDECPQCMATTTTACGVCRRR-----EPPVYR 83

QY 99 GKMRPSPG--PDRFCIPDRYRAQVOLL-CPGGEAPRARKVRLVASCCKR 146
DB 84 --YKSLGPPPSACTYGAIRYERMAWMCPIGSDPRV-LLPVALSCRCAR 131

RESULT 4
T22025
hypothetical protein F40E10.4 - Caenorhabditis elegans (fragment)
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C/Accession: T22025
R/Smyle, R.
submitted to the EMBL Data Library, February 1996
A/Reference number: Z19503
A/Accession: T22025
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-601 <WLL>
A/Cross-references: EMBL:Z69792; PIDN:CAA93668.1; GSPDB:GN00028; CESP:F40E10.4
C/Experimental source: clone F40E10
C/Genetics:
A/Gene: CESP:F40E10.4
A/Map position: X

Query Match 8.3%; Score 87.5; DB 2; Length 601;
Best Local Similarity 20.6%; Pred. No. 3.9; Indels 19; Gaps 2;
Matches 21; Conservative 17; Mismatches 45; Indels 19; Gaps 2;

QY 44 HHPEFKDVSEVSCRELFHFRVYTDGPCRSAPVTELVCSGCCGPARLLPNAIGRKWMR 103
|||||

DB 510 HPSGHEDEKRIKCDKQFRHHIENECSVDRIKIAECNCGC----- 553

QY 104 PSGDPFRCPDRYRAQVOLL-CPGGEAPRARKVRLVASCCKR 145
DB 554 --GEONCCTAIVAKKQKRVKMTCKNG-TTKISTVHIIRQCCE 592

RESULT 5
D89711
protein F40E10.4 [imported] - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Nov-2001
C/Accession: D89711
R/Anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A/Title: Genome sequence of the nematode C. elegans: a platform for investigating biology
A/Reference number: A75000; MUID:99069613; PMID:9851916
A/Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C_elegans/
A/Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A/Accession: D89711
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-601 <STO>
A/Cross-references: GB:chr_X; PIDN:CAA93668.1; PID:93877014; GSPDB:GN00028; CESP:F40E10.4
C/Genetics:
A/Gene: F40E10.4
A/Map position: X

Query Match 8.3%; Score 87.5; DB 2; Length 601;
Best Local Similarity 20.6%; Pred. No. 3.9; Indels 19; Gaps 2;
Matches 21; Conservative 17; Mismatches 45; Indels 19; Gaps 2;

QY 44 HHPEFKDVSEVSCRELFHFRVYTDGPCRSAPVTELVCSGCCGPARLLPNAIGRKWMR 103
|||||
DB 510 HPSGHEDEKRIKCDKQFRHHIENECSVDRIKIAECNCGC----- 553

QY 104 PSGDPFRCPDRYRAQVOLL-CPGGEAPRARKVRLVASCCKR 145
DB 554 --GEONCCTAIVAKKQKRVKMTCKNG-TTKISTVHIIRQCCE 592

RESULT 6
B3665
silt protein 2 precursor - fruit fly (Drosophila melanogaster)
C/Species: Drosophila melanogaster
C/Date: 30-Apr-1991 #sequence_revision 30-Apr-1991 #text_change 02-Aug-2002
C/Accession: B3665
R/Rothberg, J.M.; Jacobs, J.R.; Goodman, C.S.; Artavanis-Tsakonas, S.
Genes Dev. 4, 2169-2187, 1990
A/Title: silt: an extracellular protein necessary for development of midline glia and con
A/Reference number: B3665; MUID:91099665; PMID:2176636
A/Accession: B3665
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-1469 <ROT>
A/Cross-references: GB:X53959
C/Genetics:
A/Gene: FlyBase:silt
A/Cross-references: FlyBase:FBgn0003425
C/Superfamily: fruit fly silt protein; EGF homology; leucine-rich alpha-2-glycoprotein re
F/66-91/Domains: proteoglycan amino-terminal homology <PAH1>
F/101-124/Domains: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>
F/125-148/Domains: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>
F/149-172/Domains: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>
F/173-196/Domains: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>
F/197-220/Domains: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>
F/228-272/Domains: proteoglycan carboxyl-terminal homology <PCSI1>
F/288-313/Domains: proteoglycan amino-terminal homology <PAH2>
F/323-346/Domains: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>
F/347-370/Domains: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>
F/371-394/Domains: leucine-rich alpha-2-glycoprotein repeat homology <LRR8>
F/395-418/Domains: leucine-rich alpha-2-glycoprotein repeat homology <LRR9>
F/419-442/Domains: leucine-rich alpha-2-glycoprotein repeat homology <LRR10>

F:450-494/Domain: proteoglycan carboxyl-terminal homology <PCS2>
 F:512-537/Domain: proteoglycan amino-terminal homology <PAH3>
 F:547-571/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR11>
 F:572-595/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR12>
 F:596-619/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR13>
 F:620-643/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR14>
 F:651-695/Domain: proteoglycan carboxyl-terminal homology <PCS3>
 F:708-733/Domain: proteoglycan amino-terminal homology <PAH4>
 F:743-766/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR15>
 F:767-790/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR16>
 F:846-890/Domain: proteoglycan carboxyl-terminal homology <PCS4>
 F:1028-1061/Domain: EGF homology <EGF2>
 F:1068-1099/Domain: EGF homology <EGF2>
 F:1115-1148/Domain: EGF homology <EGF1>

Query Match 7.9%; Score 82.5; DB 2; Length 1469;
 Best Local Similarity 20.0%; Pred. No. 27;

Matches 32; Conservative 23; Mismatches 62; Indels 43; Gaps 6;

QY 4 QAFKNDATETIIPELGEYPEPP-PELENNKTMRAENGRR-PPHAPETK-----50

DB 1333 QEEEDDEQDFMDETPHKEEVPDCE-----NKCRRGRCVPSNARJGCKCKHGR 1387

QY 51 -----DVESEYCRELHFRVYTDGCRSAKPVTELVCSGCGPARLPMALGRGKMRPSC 106

DB 1388 GRVCDQASTCRKEQVREYTYENDCRSRQPLKAKCVGGG-----1428

QY 107 PDFRCIPDRYRAQRVOLLCGGEAPRAKRVLVASCCKR 146

DB 1429 -NOCCAKIVRRRRVAVC-SNNRKYTKNDIYRKGCGTK 1466

RESULT 7

S71793

head-inducing factor Cerberus - African clawed frog

C/Species: Xenopus laevis (African clawed frog)

C/Date: 12-Feb-1998 #sequence_revision 13-Mar-1998 #text_change 21-Jul-2000

C/Accession: S71793

R:Boumester, T.; Kim, S.H.; Saeki, Y.; Lu, B.; de Robertis, E.M.

Nature 382, 595-601, 1996

A/Title: Cerberus is a head-inducing secreted factor expressed in the anterior endoderm

A/Reference number: S71793; MUID:9638220; PMID:8757128

A/Accession: S71793

A/Status: preliminary; nucleic acid sequence not shown

A/Molecule type: mRNA

A/Residues: 1-270 <BOU>

A/Cross-references: EMBL:U64831; NID:G1513087; PIND:AAC6012.1; PID:G1513088

Query Match 7.7%; Score 81; DB 2; Length 270;

Best Local Similarity 17.7%; Pred. No. 6.8;

Matches 28; Conservative 34; Mismatches 70; Indels 26; Gaps 6;

QY 4 QAFKNDATETIIPELGEYPEPPPELENNKTMRAENGRRPHHPEFTKD/SEVSCRELHPT 63

DB 122 RSDKDKNTET-----EKRGACFMNNFLVKGANGAPQNTSHSKAOELMKCKTLPT 175

QY 64 RYVTDGCRSAKPVTELVCSCGPARLPMALGRGKMRPSCDPR-----CIPRYRA 118

DB 176 QNIVHENC-DRMYIQNNLCFGKCSILHV-----PNOQDRRNTCSHCLPSKFTL 222

QY 119 QRVOLLCGGEAPRAKRVLVASCCK-RLTRFHNOSE 155

DB 223 NHLLNCTGSKNV-VKVMWVEBCTCAHNSPHQTAQ 259

RESULT 8

T12534

hypothetical protein DKFP434B094.1 - human (fragment)

C/Species: Homo sapiens (man)

C/Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 02-Sep-2000

C/Accession: T12534

R:Mamuti, R.; Heuber, D.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, June 1999

A/Reference number: 217524
 A/Accession: T12534
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-715 <MAN>
 A/Cross-references: EMBL:AL080149
 A/Experimental source: adult testis; clone DKFP434B094
 C/Genetics:
 A/Note: DKFP434B094.1
 C/Superfamily: bromodomain homology
 F:113-168/Domain: bromodomain homology <BRO>

Query Match 7.7%; Score 80.5; DB 2; Length 715;

Best Local Similarity 23.9%; Pred. No. 20;

Matches 49; Conservative 20; Mismatches 87; Indels 49; Gaps 9;

QY 2 QHAFKNDATETIIPELGEY-PEPPPELENNKTM---NRAENGRRPHHPEFTKDYSEVSC 57

DB 294 GEGFEDGALGPEGEGDGSPPKLEPSDALPLPSNSETSEBP-----TLKPELNPE 349

QY 58 RELHFRVYTDGCRSAKPVTELVCSCGPARLPMALGRGKMRPSCDPRCIPDRYR 117

DB 350 GSKLFRKRVTFDSEHSACQSLV-SG-----RPEPTTASSGDVPA 390

QY 118 AQRVOLLCGGEAPRAKRVLVASCCKRLTRFHNOSELKQGTAAAPQGRK-----171

DB 391 AASAVAEPAEDVNRRTSVLF---CKSKSVS---PPKSAKNTETQPTSPQGTGTFSLV 444

QY 172 -----PPRARSARANOAELE 187

DB 445 LPRLETLQPRRSRST-CGDSVE 468

RESULT 9

A3665

slit protein 1 precursor - fruit fly (Drosophila melanogaster)

C/Species: Drosophila melanogaster

C/Date: 30-Apr-1991 #sequence_revision 30-Apr-1991 #text_change 02-Aug-2002

C/Accession: A3665; A31640; S13523

R:Rothenberg, J.M.; Jacobs, J.R.; Goodman, C.S.; Artavanis-Tsakonas, S.

Genes Dev. 4, 2169-2187, 1990

A/Title: slit: an extracellular protein necessary for development of midline glia and con

A/Reference number: A3665; MUID:9109665; PMID:2176636

A/Accession: A31640

A/Molecule type: DNA

A/Residues: 881-1182, 'G', '1185-1404', 'GT', '1463-1464', 'YHA', 'RQ2>

A/Cross-references: GB:M23543; NID:G340939; PID:G514357

C/Genetics:
 A/Gene: FlyBase:slit

A/Cross-references: FlyBase:FBgn0003425

A/Introns: 1351/3

C/Superfamily: fruit fly slit protein; EGF homology; leucine-rich alpha-2-glycoprotein re

C/Keywords: alternative splicing; growth factor

F:66-91/Domain: proteoglycan amino-terminal homology <PAH1>

F:101-124/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR1>

F:125-148/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR2>

F:149-172/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR3>

F:173-196/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR4>

F:197-220/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR5>

F:228-272/Domain: proteoglycan carboxyl-terminal homology <PCS1>

F:288-313/Domain: proteoglycan amino-terminal homology <PAH2>

F:323-346/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR6>

F:347-370/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR7>

F:371-394/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR8>

F:395-418/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR9>
 F:419-442/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR10>
 F:450-494/Domain: proteoglycan carboxyl-terminal homology <PC8>
 F:512-537/Domain: proteoglycan amino-terminal homology <PA3>
 F:547-571/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR11>
 F:572-595/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR12>
 F:596-619/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR13>
 F:620-643/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR14>
 F:651-695/Domain: proteoglycan carboxyl-terminal homology <PC3>
 F:708-733/Domain: proteoglycan amino-terminal homology <PA4>
 F:743-766/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR15>
 F:767-790/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR16>
 F:791-814/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR17>
 F:815-838/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR18>
 F:846-890/Domain: proteoglycan carboxyl-terminal homology <PC4>
 F:1028-1061/Domain: EGF homology <EGF>
 F:1068-1099/Domain: EGF homology <EGF2>
 F:1115-1148/Domain: EGF homology <EGF1>

Query Match 7.5%; Score 80; DB 2; Length 1480;
 Best Local Similarity 18.0%; Pred. No. 45;
 Matches 31; Conservative 23; Mismatches 62; Indels 56; Gaps 5;

QY 4 QAEKNDATETIIPETGEYEP--PELEN-----NKTWN 34
 DB 1333 QEEBDDSDMDENPHKEPVPCLNKCRGRSCVPSNNDGYCKKHGRGYCD 1392
 QY 35 RAENGRRPHHPETKQVSEYSCRELFHYVTDGPRSAKPYTELVSQCQCPALLRN 94
 DB 1393 QGEESTEP-----TVTAISTCKEQVREYTYENDSRQPLKYAKCVGCG----- 1439
 QY 95 AIGRKWRRSPGDFRCIPDRYRAORVOLLCPGGEAPRAKRVLVASCCKR 146
 DB 1440 -----NCCAKIVRRKRYMVC-SNNRKYTNLDIVKCGCTK 1477

RESULT 10

PNO141

Lutropin beta chain - sperm whale

N/Alternate names: luteinizing hormone beta chain

C/Species: Phryner catodon (sperm whale)

C/Date: 07-May-1993 #sequence_revision 07-May-1993 #text_change 07-May-1999

C/Accession: PNO141

R/Pankov, Y.A.; Karasev, V.S.

Biochimica 49, 1004-1018, 1984

A/Title: Luteinizing hormone of the sperm-whale: amino acid sequence of reduced and carb

A/Reference number: PNO141; MUID:84281133; PMID:6466737

A/Accession: PNO141

A/Molecule type: protein

A/Residues: 1-118 <PAM>

A/Note: article in Russian with English abstract

C/Superfamily: pituitary glycoprotein hormone beta chain

C/Keywords: glycoprotein; hormone

F:9-34,23-57,26-88,110,72-100,90-93/Disulfide bonds: #status predicted

F:13/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 7.5%; Score 79; DB 2; Length 118;
 Best Local Similarity 27.6%; Pred. No. 45;
 Matches 34; Conservative 14; Mismatches 55; Indels 20; Gaps 7;

QY 57 CRELHFRVYVTDGPRSAKPYTELVSQCQCPARLLPNAIGRKWRRSPGDFRCIPDR 115
 DB 9 CRFINATLAONZACPCITFTTISICAGYCSMRVLPAL-----PVPZPVCYRQ 61
 QY 116 YRAORVOLLCPGGEAPRAKRVLVASCCK--RLTFHNSSELKDFGTAAAPQK-GRK 171
 DB 62 LRFASIRLPGCPGVNBMV-SFVALSCHGCPRLSS-----SDGPRAPQPLACNRS 113

QY 172 PRP 174
 DB 114 PRP 116

RESULT 11

151824

tumor-suppressive gene - rat

C/Species: Rattus sp. (rat)

C/Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 05-Nov-1999

C/Accession: 151824

R/Sakiyama, S.; Ozaki, T.; Enomoto, H.

Adv. Enzyme Regul. 34, 247-255, 1994

A/Title: Molecular cloning and characterization of a cDNA showing tumor-suppressive ac

A/Reference number: 151824; MUID:95028756; PMID:7942277

A/Accession: 151824

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-178 <RES>

A/Cross-references: GB:672637; NID:9619290; PIDN:AAB32215.1; PID:9619291

Query Match 7.5%; Score 79; DB 2; Length 178;
 Best Local Similarity 25.0%; Pred. No. 68;
 Matches 38; Conservative 18; Mismatches 68; Indels 28; Gaps 9;

QY 57 CRELHFRVYVTDGPRSAKPYTELVSQCQCPARLLPNAIGRKWRRSPGDF-----RC 111
 DB 34 CEAKITQIVGHSGC-EAKSIQNACUGQCF-SYVNTF-----POSTESLVHCDSC 84
 QY 112 IPDRYRAORVOLLCPG-GEAPRAK-RVLVASCCKRLTRFHNSSELKDF--GTEAAPQ 167
 DB 85 MPAQSMWEIVTLCEGHEVPRVDKLVKIVHSCQACGKEPSHGLNVYMGEDGGSQ 144
 QY 168 KG---RKPRPARSAK-----ANQAELENA 189
 DB 145 PSHSHSHPHGCGQTPPEEPGAPQVEEGA 176

RESULT 12

A47291

Probable finger protein - rat

C/Species: Rattus norvegicus (Norway rat)

C/Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 05-Nov-1999

C/Accession: A47291

R/Ozaki, T.; Sakiyama, S.

Proc. Natl. Acad. Sci. U.S.A. 90, 2593-2597, 1993

A/Title: Molecular cloning and characterization of a cDNA showing negative regulation in

A/Reference number: A47291; MUID:93219329; PMID:8385338

A/Accession: A47291

A/Status: preliminary

A/Molecule type: nucleic acid

A/Residues: 1-178 <ODA>

A/Cross-references: GB:X66872; NID:9296927; PIDN:CAA47344.1; PID:9296928

A/Experimental source: fibroblast 3Y1 cells

A/Note: sequence extracted from NCBI backbone (NCBIN:128602, NCBI:P:128603)

C/Keywords: DNA binding; zinc finger

Query Match 7.5%; Score 79; DB 2; Length 178;
 Best Local Similarity 25.0%; Pred. No. 68;
 Matches 38; Conservative 18; Mismatches 68; Indels 28; Gaps 9;

QY 57 CRELHFRVYVTDGPRSAKPYTELVSQCQCPARLLPNAIGRKWRRSPGDF-----RC 111
 DB 34 CEAKITQIVGHSGC-EAKSIQNACUGQCF-SYVNTF-----POSTESLVHCDSC 84
 QY 112 IPDRYRAORVOLLCPG-GEAPRAK-RVLVASCCKRLTRFHNSSELKDF--GTEAAPQ 167
 DB 85 MPAQSMWEIVTLCEGHEVPRVDKLVKIVHSCQACGKEPSHGLNVYMGEDGGSQ 144

QY 168 KG---RKPRPARSAK-----ANQAELENA 189
 DB 145 PSHSHSHPHGCGQTPPEEPGAPQVEEGA 176

RESULT 13

T29197

hypothetical protein T03f1.9 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

•
•
•
•

GenCore version 5.1.4_p5 4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 28, 2003, 14:11:46 ; Search time 14.6933 Seconds
(without alignments)
536.332 Million cell updates/sec

Title: US-09-867-274-2

Perfect score: 1049

Sequence: 1 OGMQAFKNDATETIPELGEY.....KPRPRASAKNOALEENAY 190

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1049	100.0	213	1	SOST_HUMAN
2	1033	98.5	213	1	SOST_CERAE
3	974	92.9	213	1	SOST_RAT
4	944	90.0	211	1	SOST_MOUSE
5	901.5	85.9	176	1	SOST_BOVIN
6	95.5	9.1	493	1	AR12_HUMAN
7	94.5	9.0	159	1	LSHB_MELGA
8	89.5	8.5	492	1	AR12_MOUSE
9	87	8.3	958	1	HIG_DROME
10	85	8.1	180	1	DAN_HUMAN
11	83.5	8.0	141	1	LSHB_TRIUV
12	81.5	7.8	670	1	BCAL_HUMAN
13	80	7.6	720	1	CTSE_HUMAN
14	80	7.6	1480	1	SLIT_DROME
15	79.5	7.6	318	1	NK2E_MOUSE
16	79	7.5	118	1	LSHB_PHYCA
17	79	7.5	178	1	DAN_RAT
18	79	7.5	355	1	KLF2_HUMAN
19	78.5	7.5	166	1	LSHB_COTJA
20	78.5	7.5	969	1	STL_MYCTU
21	78	7.4	1210	1	EGFR_HUMAN
22	77.5	7.4	951	1	SFR8_HUMAN
23	77	7.3	118	1	LSHB_BALAC
24	77	7.3	510	1	IAL_HUMAN
25	76.5	7.3	220	1	NOL3_MOUSE
26	76.5	7.3	1253	1	SHK2_HUMAN
27	76	7.2	398	1	MUB1_XENLA
28	76	7.2	1213	1	JAG1_BRERA
29	76	7.2	2142	1	BAT2_HUMAN
30	75.5	7.2	318	1	NK2E_RAT
31	75.5	7.2	1436	1	MC11_BOVIN
32	75	7.1	141	1	LSHB_CERST
33	74.5	7.1	972	1	STL_MYCLE

34	74	7.1	138	1	LSHB_MACRU
35	74	7.1	269	1	YAS3_ARCFU
36	74	7.1	470	1	KLF4_HUMAN
37	74	7.1	474	1	KLF4_MOUSE
38	74	7.1	544	1	H15_DROME
39	74	7.1	702	1	HELS_AERPE
40	74	7.1	877	1	INCE_CHICK
41	74	7.1	2715	1	TRX2_HUMAN
42	73.5	7.0	682	1	SNK_MOUSE
43	73.5	7.0	682	1	SNK_RAT
44	73.5	7.0	685	1	SNK_HUMAN
45	73	7.0	178	1	DAN_MOUSE

ALIGNMENTS

RESULT 1

ID	SOST_HUMAN	STANDARD;	PRT;	213 AA.
AC	O9BOB4;			
DT	15-JUN-2002 (Rel. 41, Created)			
DT	15-JUN-2002 (Rel. 41, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Sclerostin precursor.			
GN	SOST.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	Pubmed=1181578;			
RA	Balemans W., Ebeling M., Patel N., van Hul E., Olson P., Dioszegi M.,			
RA	Lacca C., Wuyts W., van den Ende J., Willems P., Paes-Alves A.F.,			
RA	Hill S., Bueno M., Ramos F.J., Taccioni P., Dikkers F.G., Stratakis C.,			
RA	Lindpaintner K., Vickery B., Foerzner D., Van Hul W.;			
RT	"Increased bone density in sclerosteosis is due to the deficiency of a			
RT	novel secreted protein (SOST).";			
RL	Hum. Mol. Genet. 10:537-543(2001).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	Pubmed=1179006;			
RA	Brunkow M.E., Gardner J.C., Van Ness J., Paepers B.W., Kovacevich B.R.,			
RA	Proll S., Skonier J.E., Zhao L., Sabo P.J., Fu Y.H., Alisch R.S.,			
RA	Gillet L., Colbert T., Taccioni P., Galas D., Hamersma H.,			
RA	Beighton P., Mulligan J.T.;			
RT	"Bone dysplasia sclerosteosis results from loss of the SOST gene			
RT	product, a novel cystine knot-containing protein.";			
RL	Am. J. Hum. Genet. 68:577-589(2001).			
CC	- FUNCTION: Seems to play a role in bone homeostasis.			
CC	- SUBCELLULAR LOCATION: Secreted (Potential).			
CC	- TISSUE SPECIFICITY: Widely expressed at low levels with highest			
CC	levels in bone, cartilage, kidney, liver, bone marrow and primary			
CC	osteoblasts differentiated for 21 days.			
CC	- DISEASE: Defects in SOST are the cause of sclerosteosis, a			
CC	progressive sclerosing bone dysplasia.			
CC	- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTINE KNOT-LIKE (CTCK) DOMAIN.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.ebi.ac.uk/announcements			
CC	or send an email to license@ebi.ac.uk).			
CC	-----			
DR	EMBL; AF331844; AAK16158.1; -			
DR	EMBL; AF326736; AAK13451.1; -			
DR	EMBL; AF326739; AAK13454.1; -			
DR	GeneW; HGNC:13771; SOST.			
DR	MIW; 605740; -			
DR	MIW; 269500; -			

DR InterPro; IPR000359; Cys_knot.
 DR SMART; SM00041; CT; 1.
 DR PROSITE; PS01185; CTCK_1; FALSE_NEG.
 DR PROSITE; PS01225; CTCK_2; FALSE_NEG.
 KW Signal; Glycoprotein.
 FT SIGNAL 1 23 POTENTIAL.
 FT CHAIN 24 213 SCLEROSTIN.
 FT DOMAIN 82 172 CTCK.
 FT CARBOHYD 53 53 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 175 175 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 213 AA; 24030 MW; 30DBD5CE73D5B82 CRC64;
 Query Match 100.0%; Score 1049; DB 1; Length 213;
 Best Local Similarity 100.0%; Pred. No. 5e-88;
 Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QGWAQKNDATETIIPELGEYEPPEPELENNKTMNRAENGSRPHHPETKDVSEYSREL 60
 DB 24 QGWAQKNDATETIIPELGEYEPPEPELENNKTMNRAENGSRPHHPETKDVSEYSREL 83
 QY 61 HFTRYVTDGPRSAKPYTELVCSCGCGPARLLPNAIGRGMWRSPGDFRCIPDRYRAQR 120
 DB 84 HFTRYVTDGPRSAKPYTELVCSCGCGPARLLPNAIGRGMWRSPGDFRCIPDRYRAQR 143
 QY 121 VQLCPGGEAPRARKVRLVASCKCKRLTRFHNQSELKDFGEARPOKGRKPRPARSAK 180
 DB 144 VQLCPGGEAPRARKVRLVASCKCKRLTRFHNQSELKDFGEARPOKGRKPRPARSAK 203
 QY 181 ANQAELENAY 190
 DB 204 ANQAELENAY 213
 RESULT 2
 SOST_CERAE STANDARD; PRT; 213 AA.
 ID SOST_CERAE
 AC Q9BG78;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Sclerostin precursor.
 GN SOST.
 OS Cercopithecus aethiops (Green monkey) (Grivet).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecinae; Cercopithecus.
 NC NCB1_TaxID=9534;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21090529; PubMed=11179006;
 RA Brunkow M.E., Gardner J.C., Van Ness J., Paepers B.W., Kovacevich B.R.,
 RA Proll S., Skonier J.E., Zhao L., Sabo P.J., Fu Y.H., Alish R.S.,
 RA Gillett L., Colbert T., Tacconi P., Galas D., Hamerema H.,
 RA Beighton P., Mulligan J.T.;
 RT "Bone dysplasia sclerosteosis results from loss of the SOST gene
 product, a novel cysteine knot-containing protein.";
 RL Am. J. Hum. Genet. 68:577-589(2001).
 CC -!- FUNCTION: Seems to play a role in bone homeostasis (By
 similarity).
 CC -!- SUBCELLULAR LOCATION: Secreted (Potential).
 CC -!- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTINE KNOT-LIKE (CTCK) DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF326742; AAK13457.1; -
 DR InterPro; IPR000359; Cys_knot.
 DR PROSITE; PS01185; CTCK_1; FALSE_NEG.
 DR PROSITE; PS01225; CTCK_2; FALSE_NEG.
 KW Signal; Glycoprotein.

DR PROSITE; PS01185; CTCK_1; FALSE_NEG.
 DR PROSITE; PS01225; CTCK_2; FALSE_NEG.
 KW Signal; Glycoprotein.
 FT SIGNAL 1 23 POTENTIAL.
 FT CHAIN 24 213 SCLEROSTIN.
 FT DOMAIN 82 172 CTCK.
 FT CARBOHYD 53 53 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 175 175 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 213 AA; 23908 MW; 6DA7B5DE674728A CRC64;
 Query Match 98.5%; Score 1033; DB 1; Length 213;
 Best Local Similarity 98.4%; Pred. No. 1.4e-86;
 Matches 187; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 QGWAQKNDATETIIPELGEYEPPEPELENNKTMNRAENGSRPHHPETKDVSEYSREL 60
 DB 24 QGWAQKNDATETIIPELGEYEPPEPELENNKTMNRAENGSRPHHPETKDVSEYSREL 83
 QY 61 HFTRYVTDGPRSAKPYTELVCSCGCGPARLLPNAIGRGMWRSPGDFRCIPDRYRAQR 120
 DB 84 HFTRYVTDGPRSAKPYTELVCSCGCGPARLLPNAIGRGMWRSPGDFRCIPDRYRAQR 143
 QY 121 VQLCPGGEAPRARKVRLVASCKCKRLTRFHNQSELKDFGEARPOKGRKPRPARSAK 180
 DB 144 VQLCPGGEAPRARKVRLVASCKCKRLTRFHNQSELKDFGEARPOKGRKPRPARSAK 203
 QY 181 ANQAELENAY 190
 DB 204 ANQAELENAY 213
 RESULT 3
 SOST_RAT STANDARD; PRT; 213 AA.
 ID SOST_RAT
 AC Q99P67;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Sclerostin precursor.
 GN SOST.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 NC NCB1_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley;
 RX MEDLINE=21090529; PubMed=11179006;
 RA Brunkow M.E., Gardner J.C., Van Ness J., Paepers B.W., Kovacevich B.R.,
 RA Proll S., Skonier J.E., Zhao L., Sabo P.J., Fu Y.H., Alish R.S.,
 RA Gillett L., Colbert T., Tacconi P., Galas D., Hamerema H.,
 RA Beighton P., Mulligan J.T.;
 RT "Bone dysplasia sclerosteosis results from loss of the SOST gene
 product, a novel cysteine knot-containing protein.";
 RL Am. J. Hum. Genet. 68:577-589(2001).
 CC -!- FUNCTION: Seems to play a role in bone homeostasis (By
 similarity).
 CC -!- SUBCELLULAR LOCATION: Secreted (Potential).
 CC -!- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTINE KNOT-LIKE (CTCK) DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF326741; AAK13456.1; -
 DR InterPro; IPR000359; Cys_knot.
 DR PROSITE; PS01185; CTCK_1; FALSE_NEG.
 DR PROSITE; PS01225; CTCK_2; FALSE_NEG.
 KW Signal; Glycoprotein.

FT SIGNAL 1 28 POTENTIAL.
 FT CHAIN 29 213 SCLEROSTIN.
 FT DOMAIN 82 172 CTCK.
 FT CARBOHYD 53 53 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 175 175 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 213 AA; 23974 MW; 6C56C878C8D684B CRC64;

Query Match 92.9%; Score 974; DB 1; Length 213;
 Best Local Similarity 92.1%; Pred. No. 3e-81; Mismatches 8; Indels 0; Gaps 0;
 Matches 175; Conservative 7;

QY 1 OGQWAFKNDATETIIPELGEPPEPELENNKTNRANGRPHPHPEFYKDVSEYSCREL 60
 DB 24 OGQWAFKNDATETIIPGLREYEPPEPELENNQTNRAENGRRPHHPHYDKDVSEYSCREL 83
 QY 61 HFTRYVDGPRSAKPVTELVCSGCCGPARLLPNAIGRKWMRPSGDPFRICIDRYRAOR 120
 DB 84 HFTRYVDGPRSAKPVTELVCSGCCGPARLLPNAIGRKWMRPSGDPFRICIDRYRAOR 143
 QY 121 VOLCPGGEAPRARKVRLVASCCKRLTRFNHNSLKDGTENARPQKGRKPRPARSAK 180
 DB 144 VOLCPGGEAPRARKVRLVASCCKRLTRFNHNSLKDGTENARPQKGRKPRPARSAK 203
 QY 181 ANQAELENAY 190
 DB 204 ANQAELENAY 213

RESULT 4
 SOST_MOUSE STANDARD; PRT; 211 AA.
 ID SOST_MOUSE Q99P68; Q99D17;
 AC 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Sclerostin precursor.
 GN SOST.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129/SvJ;
 RX MEDLINE=21090529; PubMed=11179006;
 RA Brunkow M.E., Gardner J.C., Van Ness J., Paepers B.W., Kovacevich B.R.,
 RA Prohl S., Skonier J.E., Zhao L., Sabo P.J., Fu Y.H., Alisch R.S.,
 RA Gillett L., Colbert T., Tacconi P., Galas D., Hamersma H.,
 RA Beighton P., Mulligan J.T.;
 RA "bone dysplasia sclerosteosis results from loss of the SOST gene
 RT product, a novel cysteine knot-containing protein.";
 RL Am. J. Hum. Genet. 68:577-589(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Head;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi U., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,
 RA Saito T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Glassi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsumoto Y., Niki K., Pesole G., Quackenbush J.,
 RA Sakai K., Staudt L., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Saki K., Okada T., Furuno M., Aono H., Baladrelli R., Batsch G.,
 RA Blake J., Botteilli D., Boujunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Guernicchi S., Hill D., Hofmann M., Hume D.A., Kamuya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombert P.,
 RA Norone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Saeki H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohatsu S.,

RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 CC -1- FUNCTION: Seems to play a role in bone homeostasis (By
 CC similarity).
 CC -1- SUBCELLULAR LOCATION: Secreted (Potential).
 CC -1- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTEINE KNOT-LIKE (CTCK) DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; AF326740; AK13455.1; -;
 DR EMBL; AK017295; BAB30678.1; -;
 DR EMBL; AF326737; AK13452.1; -;
 DR MGI; MGI:1921749; Sost.
 DR InterPro; IPR000359; Cys knot.
 DR PROSITE; PS01185; CTCK_1; FALSE NEG.
 DR PROSITE; PS01225; CTCK_2; FALSE NEG.
 KW Signal; Glycoprotein.
 FT SIGNAL 1 23 POTENTIAL.
 FT CHAIN 24 211 SCLEROSTIN.
 FT DOMAIN 80 170 CTCK.
 FT CARBOHYD 51 51 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 173 173 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 72 72 G -> D (IN REF. 2).
 SQ SEQUENCE 211 AA; 23443 MW; AEB094E358E34961 CRC64;

Query Match 90.0%; Score 944; DB 1; Length 211;
 Best Local Similarity 88.9%; Pred. No. 1.5e-78; Mismatches 9; Indels 2; Gaps 1;
 Matches 169; Conservative 10;

QY 1 OGQWAFKNDATETIIPELGEPPEPELENNKTNRANGRPHPHPEFYKDVSEYSCREL 60
 DB 24 OGQWAFKNDATETIIPGLREYEPPEPELENNQTNRAENGRRPHHPHYDKDVSEYSCREL 81
 QY 61 HFTRYVDGPRSAKPVTELVCSGCCGPARLLPNAIGRKWMRPSGDPFRICIDRYRAOR 120
 DB 82 HFTRYVDGPRSAKPVTELVCSGCCGPARLLPNAIGRKWMRPSGDPFRICIDRYRAOR 141
 QY 121 VOLCPGGEAPRARKVRLVASCCKRLTRFNHNSLKDGTENARPQKGRKPRPARSAK 180
 DB 142 VOLCPGGEAPRARKVRLVASCCKRLTRFNHNSLKDGTENARPQKGRKPRPARSAK 201
 QY 181 ANQAELENAY 190
 DB 202 ANQAELENAY 211

RESULT 5
 SOST_BOVIN STANDARD; PRT; 176 AA.
 ID SOST_BOVIN Q9BQ79;
 AC 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Sclerostin (Fragment).
 GN SOST.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OC NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=21090529; PubMed=11179006;
 RA Brunkow M.E., Gardner J.C., Van Ness J., Paepers B.W., Kovacevich B.R.,
 RA Prohl S., Skonier J.E., Zhao L., Sabo P.J., Fu Y.H., Alisch R.S.,
 RA Gillett L., Colbert T., Tacconi P., Galas D., Hamersma H.,

RA Beighton P., Mulligan J.T.;
 RT "bone dysplasia sclerosteosis results from loss of the SOST gene
 product: a novel cysteine knot-containing protein.";
 RL Am. J. Hum. Genet. 68:577-589(2001).
 CC -1- FUNCTION: Seems to play a role in bone homeostasis (By
 similarity).
 CC -1- SUBCELLULAR LOCATION: Secreted (Potential).
 CC -1- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTEINE KNOT-LIKE (CTCK) DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF326738; AAK13453.1; -;
 DR InterPro; IPR000359; Cys_knot.
 DR SMART; SM00041; CT; 1.
 DR PROSITE; PS01185; CTCK_1; FALSE_NEG.
 DR PROSITE; PS01225; CTCK_2; FALSE_NEG.
 KW Glycoprotein.
 FT DOMAIN 1 141 CTCK.
 FT CARBOHYD 22 22 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 144 144 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT NON_TER 176 176
 SQ SEQUENCE 176 AA, 19743 MW, 35FOCA61A425F4DB CRC64;
 Query Match 85.9%; Score 901.5; DB 1; Length 176;
 Best Local Similarity 93.8%; Pred. No. 8.9e-75;
 Matches 166; Conservative 2; Mismatches 8; Indels 1; Gaps 1;
 QY 8 NDATETIIPELGEYEPPEPELENNKTKMRAENGRRPHHPETKDVSYSCRELIHFTRYVT 67
 DB 1 NDATETIIPELGEYEPPEPELENNKTKMRAENGRRPHHPETKDVSYSCRELIHFTRYVT 59
 QY 68 DGPGRSAKPVTELVCSGCGCPARLLPNAIGKWKWSPSGDFRCIPRYVYAOGVOLCPG 127
 DB 60 DGPGRSAKPVTELVCSGCGCPARLLPNAIGKWKWSPSGDFRCIPRYVYAOGVOLCPG 119
 QY 128 GEAPRRARKVRLVASCKCKRLTRFRNOSLKFCTEARPOKGRKPPRARSANQQA 184
 DB 120 GAAPRRARKVRLVASCKCKRLTRFRNOSLKFCTEARPOKGRKPPRARSANQQA 176
 RESULT 6
 ARI2 HUMAN STANDARD; PRT; 493 AA.
 ID ARI2 HUMAN STANDARD; PRT; 493 AA.
 AC 095376; O95B26;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Atrialn-2 protein homolog (ARI-2) (Triadi protein) (HT005).
 GN ARI2 OR ARI2 OR TRIAD1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OC NCBI_TaxID=9606;
 OX 1
 RN
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9349709; Pubmed=10422847;
 RA van der Reijden B.A., Erpelinck-Verschueren C.A.J., Loewenberg B.,
 RA Jansen J.H.;
 RT "RIRDS: a new class of proteins with a novel cysteine-rich
 RT signature";
 RL Protein Sci. 8:1557-1561(1999).
 RN
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20341325; Pubmed=10880484;
 RA Aguilera M., Oliveros M., Martinez-Padron M., Barbas J.A., Ferrus A.;
 RA "Atrialn-1: a vital Drosophila gene is required in development and

RT defines a new conserved family of ring-finger proteins.";
 RL Genetics 155:1231-1244(2000).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX TISUB=Hypothalamus;
 RX MEDLINE=20402571; Pubmed=10931946;
 RA Hu R.-M., Han Z.-G., Song H.-D., Peng Y.-D., Huang Q.-H., Ren S.-X.,
 RA Gu Y.-Y., Huang C.-H., Li Y.-B., Jiang C.-L., Fu G., Zhang Q.-H.,
 RA Gu B.-W., Dai M., Mao Y.-F., Gao G.-F., Rong R., Ye M., Zhou J.,
 RA Xu S.-H., Gu J., Shi J.-X., Jin W.-R., Zhang C.-K., Wu T.-M.,
 RA Huang G.-Y., Chen Z., Chen M.-D., Chen J.-L.;
 RT "Gene expression profiling in the human hypothalamus-pituitary-adrenal
 RT axis and full-length cDNA cloning";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:9543-9548(2000).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC TISUB=Muscle;
 RA Strausberg R.;
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: MIGHT ACT AS AN E3 UBIQUITIN-PROTEIN LIGASE, OR AS PART
 CC OF E3 COMPLEX, WHICH ACCEPTS UBIQUITIN FROM SPECIFIC E2 UBIQUITIN-
 CC CONJUGATING ENZYMES, SUCH AS UBE2L3/UBC4, AND THEN TRANSFERS IT
 CC TO SUBSTRATES.
 CC -1- SUBUNIT: INTERACTS WITH UBE2L3 (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- SIMILARITY: CONTAINS 2 RING-TYPE ZINC FINGERS.
 CC -1- SIMILARITY: CONTAINS 1 IRR-TYPE ZINC FINGER.
 CC -1- CAUTION: REP.3 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO
 CC FRAMESHIFTS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF099149; AAC82469.1; -;
 DR EMBL; AJ130978; CAA10276.1; -;
 DR EMBL; AF183427; AAG09696.1; ALT_FRAME.
 DR EMBL; BC000422; AAH00422.1; -;
 DR Genew; HGNC:690; ARI2.
 DR MIM; 605615; -;
 DR InterPro; IPR002867; Znf_GCHC.
 DR InterPro; IPR001879; Znf_CCHC.
 DR InterPro; IPR001841; Znf_Fing.
 DR Pfam; PF01485; IRR; 2.
 DR SMART; SM0184; RING; 2.
 DR SMART; SM00343; Znf_C2HC; 1.
 DR PROSITE; PS00518; ZF_RING_1; 1.
 DR PROSITE; PS00889; ZF_RING_2; 2.
 KW Ubl conjugation pathway; Nuclear protein; Coiled coil; Zinc-finger;
 KW Repeat.
 FT DOMAIN 4 75 ASP/GLU-RICH (ACIDIC).
 FT DOMAIN 22 29 POLY-GLU.
 FT ZN_FING 139 192 RING-TYPE 1.
 FT ZN_FING 208 270 IRR-TYPE.
 FT ZN_FING 300 344 RING-TYPE 2.
 FT DOMAIN 369 400 COILED COIL (POTENTIAL).
 FT DOMAIN 439 492 COILED COIL (POTENTIAL).
 FT CONFLICT 280 281 CA -> LQ (IN REF. 2).
 SQ SEQUENCE 493 AA, 57818 MW, 30AFPD327B51013 CRC64;
 Query Match 9.1%; Score 95.5; DB 1; Length 493;
 Best Local Similarity 24.2%; Pred. No. 0.2;
 Matches 45; Conservative 24; Mismatches 54; Indels 63; Gaps 11;
 QY 3 WQA-----FKNDATETIIPELGEYEPPEPELENNKTKMRAENGRRPHHPH-----PFETKD 51
 DB 100 WQVSEILIRYKSNLSQQLVERAVQNPESKHVPTS-----HPPHCAVCMQFVRKE 149
 QY 52 -----VSEY-----SCRELIHFTRYVTDGPGRSAKPVTELVCSGCGCPAR-----LLPN 94


```

Db 150 NLLSLACOHQFCRSCWBOHCSVLXKXG-----VGVVSCMAQDCPLRPEDFVPLLPN 203
Qy 95 AIGRGKMMRSGPDRFCIPDRYRAORVOLLCPGGEA-----PRARKYRL-----VAS 141
Db 204 BELREKRRRLFRDY--VESHYOLQ-----LCPADCPMTIRVOEPARRVOCNRCNEVEFC 257
Qy 142 CKCKRL 147
Db 258 FKCRQM 263

RESULT 7
LSHB_MELGA STANDARD; PRT; 159 AA.
ID ID LSHB_MELGA STANDARD; PRT; 159 AA.
AC P45646;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Lutropin beta chain precursor (Luteinizing hormone beta subunit) (LSH-beta) (LSH-B) (LH-B).
GN LHB.
OS Meleagris gallopavo (Common turkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Meleagrididae; Meleagris.
OX NCBI_TaxId=9103;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pituitary;
RX MEDLINE=95290073; PubMed=7772235;
RA You S., Foster L.K., Silsby J.L., el Halawani M.E., Foster D.N.;
RT "Sequence analysis of the turkey LH beta subunit and its regulation by gonadotrophin-releasing hormone and prolactin in cultured pituitary cells."
RL J. Mol. Endocrinol. 14:117-129(1995).
RN [2]
CC -1- FUNCTION: PROMOTES SPERMATOGENESIS AND OVULATION BY STIMULATING THE TESTES AND OVARIES TO SYNTHESIZE STEROIDS.
CC -1- SUBUNIT: HETERODIMER OF A COMMON ALPHA CHAIN AND A UNIQUE BETA CHAIN WHICH CONFERES BIOLOGICAL SPECIFICITY TO THYROTROPIN, LUTROPIN, FOLLITROPIN AND GONADOTROPIN.
CC -1- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES BETA CHAIN FAMILY.
CC CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL, L35519; AAA74125.1; ALT_INIT.
DR HSSP; P01233; 1XUL.
DR InterPro; IPR000359; Cys_knot.
DR InterPro; IPR002400; GF_cysknot.
DR InterPro; IPR001545; Gly_hormoneB.
DR Pfam; PF00007; Cys_knot_1.
DR PRINTS; PR00438; GFCYSKNOT.
DR SMART; SM00068; GHb; 1.
DR PROSITE; PS00261; GLYCO_HORMONE_BETA_1; 1.
DR PROSITE; PS00689; GLYCO_HORMONE_BETA_2; 1.
KW Hormone; Signal; Glycoprotein.
FT SIGNAL 1 39
FT CHAIN 40 159
FT DISULFID 48 96
FT DISULFID 62 111
FT DISULFID 65 149
FT DISULFID 73 127
FT DISULFID 77 129
FT DISULFID 132 139
FT CARBOHYD 52 52
FT SEQUENCE 159 AA; 16285 MW; 52B50C8C879653C6 CRC64;

```

```

Query Match 9.0%; Score 94.5; DB 1; Length 159;
Best Local Similarity 29.4%; Pred. No. 0.077;
Matches 32; Conservative 7; Mismatches 49; Indels 21; Gaps 5;

Qy 39 GGRPHHPEFTKDVSEYSCRELHPTRYTDPGCRSAKVTBLVSCGGCPARLPLMAIGR 98
Db 43 GGRPP-----CRPTNTVAVEKDECPQCMATVTTACGCGYCTR---EPVYR 85
Qy 99 GKMRPSPGDRFCIPDRYRAORVOLLCPGGEAPRRARKYRLVASCCKR 146
Db 86 SPLGRP--PQSSCTYGALRYERKALMGCPISDPRV-LLPVALSRCAR 131

RESULT 8
AR12_MOUSE STANDARD; PRT; 492 AA.
ID ID AR12_MOUSE STANDARD; PRT; 492 AA.
AC Q92AK6;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ariadne-2 protein homolog (Ari-2) (Triad1 protein) (UbcM4-interacting protein 48).
GN ARH2 OR AR12 OR TRIAD1 OR UTP48.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BA1B/c; TISSUE=Liver;
RX MEDLINE=99358765; PubMed=10431818;
RA Martinez-Noel G., Niedenthal R., Tamura T., Harbers K.;
RT "A family of structurally related RING finger proteins interacts specifically with the ubiquitin-conjugating enzyme UbcM4."
RL FEBS Lett. 454:257-261(1999).
RN [2]
CC SEQUENCE FROM N.A.
RC TISSUE=Brain, and Embryo;
RX MEDLINE=20341325; PubMed=10880484;
RA Aguilera M., Oliveros M., Martinez-Padron M., Barbás J.A., Ferrus A.;
RT "Ariadne-1: a vital Drosophila gene is required in development and defines a new conserved family of ring-finger proteins."
RL Genetics 155:1231-1244(2000).
CC -1- FUNCTION: MIGHT ACT AS AN E3 UBIQUITIN-PROTEIN LIGASE, OR AS PART OF E3 COMPLEX, WHICH ACCEPTS UBIQUITIN FROM SPECIFIC E2 UBIQUITIN-CONJUGATING ENZYMES, SUCH AS UBE2L3/UBC4, AND THEN TRANSFERS IT TO SUBSTRATES.
CC -1- SUBUNIT: INTERACTS WITH UBE2L3.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: CONTAINS 2 RING-TYPE ZINC FINGERS.
CC -1- SIMILARITY: CONTAINS 1 IRR-TYPE ZINC FINGER.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL, AF124664; AAD24573.1; -.
DR EMBL, AJ130975; CA110273.1; -.
DR MGI; MGI:1344361; Arih2.
DR InterPro; IPR002867; Znf_CGHC.
DR InterPro; IPR001878; Znf_CCHC.
DR InterPro; IPR001841; Znf_ring.
DR Pfam; PF01485; IRR; 2.
DR SMART; SM00184; RING; 2.
DR SMART; SM00343; ZNF_C2HC; 1.
DR PROSITE; PS00518; ZF_RING_1; 1.
DR PROSITE; PS00589; ZF_RING_2; 2.
KW Ubiquitination pathway; Nuclear protein; Coiled coil; Zinc-finger; Repeat.

```

FT DOMAIN 4 74 ASP/GLU-RICH (ACIDIC).
 FT DOMAIN 22 29 POLY-GLU.
 FT ZN FING 138 191 RING-TYPE 1.
 FT DOMAIN 207 269 IIR-TYPE.
 FT ZN FING 299 343 RING-TYPE 2.
 FT DOMAIN 368 399 COILED COIL (POTENTIAL).
 FT DOMAIN 438 491 COILED COIL (POTENTIAL).
 SQ SEQUENCE 492 AA; 57696 MW; B55EA54FE8C3ADAF CRC64;
 Query Match 8.5%; Score 89.5; DB 1; Length 492;
 Best Local Similarity 23.1%; Pred. No. 0.7; 55; Indels 63; Gaps 11;
 Matches 43; Conservative 23; Mismatches 55; Indels 63; Gaps 11;
 QY 3 WQA-----FKNDATETIPELGEYEPPEPELENNKTMRAENGSRPPHH-----PEETKD 51
 DB 99 WQVEIILDRYNSNAQLLVEARVQPNPSKIVPT-----AHPPHCACVCMQFYRKE 148
 QY 52 -----VSEY-----SCRELFHFRVYTDGPCRSKAKFVTELVCSGCCGPAR-----LLPN 94
 DB 149 NLLSLACOHCFRCSCWEHCSEVLVKDG-----VGVGISCAQDCPLRTDEDFVFLPN 202
 QY 95 AIGRKMWRPSPDPFRCPDPRYRAORVQLCPGGEA-----PRARKYRL-----VAS 141
 DB 203 EELDKTRRYFRDY--VESHFQLQ-----LCPGADCPMVRVQEPARRRVQCNRCSEVFC 256
 QY 142 CKCRRL 147
 DB 257 FKCRQM 262
 RESULT 9
 HIG_DROME STANDARD; PRT; 958 AA.
 ID 009101;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Locomotion-related protein HIKARU genki precursor.
 GN HIG.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RP SEQUENCE FROM N.A.
 RC TISSUE=Head;
 RX MEDLINE=93213498; Pubmed=8461133;
 RA Hoshino M., Matsuzaki F., Nabeshima Y.-I., Hama C.;
 RT "Hikaru genki, a CNS-specific gene identified by abnormal locomotion
 in Drosophila, encodes a novel type of protein.",
 RL Neuron 10:395-407(1993).
 CC - FUNCTION: HAS A ROLE IN THE DEVELOPMENT OF CNS FUNCTIONS INVOLVED
 IN LOCOMOTOR ACTIVITY.
 CC - SUBCELLULAR LOCATION: Extracellular.
 CC - ALTERNATIVE PRODUCTS: 4 ISOFORMS; 1, 2, 3 (SHOWN HERE) AND 4; ARE
 PRODUCED BY ALTERNATIVE SPLICING.
 CC - TISSUE SPECIFICITY: EXPRESSED IN PCC NEURONS AND NEUROBLASTS
 IN THE PROCEPHALIC NEUROGENIC REGION IN THE CENTRAL NERVOUS
 SYSTEM.
 CC - DEVELOPMENTAL STAGE: MOST ABUNDANT DURING AND/OR AFTER NEURONAL
 DIFFERENTIATION AND DURING CELL SPECIFICATION OR APOGEMESIS.
 CC - SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
 CC - SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
 CC - SIMILARITY: CONTAINS 4 SUSHI (SCR) DOMAINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL; D13884; BAA02984.1; -
 DR EMBL; D13885; BAA02985.1; -
 DR EMBL; D13886; BAA02986.1; -
 DR EMBL; D13887; BAA02987.1; -
 DR HSSP; P10998; IVD.
 DR FLYBASE; FBgn0010114; hig.
 DR InterPro; IPR003599; IG.
 DR InterPro; IPR003006; IG_MHC.
 DR InterPro; IPR000436; Sushi_SCR_CCP.
 DR Pfam; PF00047; ig.1.
 DR Pfam; PF00084; sushi.5.
 DR SMART; SM00032; CCP.5.
 DR SMART; SM00409; IG.1.
 DR GlycoProtein; Alternative splicing; Immunoglobulin domain; Repeat;
 KW Sushi, signal.
 FT SIGNAL 1 31
 FT CHAIN 32 958
 FT DOMAIN 630 709
 FT DOMAIN 713 769
 FT DOMAIN 772 828
 FT DOMAIN 831 891
 FT DOMAIN 893 952
 FT SITE 318 320
 FT DISULFID 714 755
 FT DISULFID 741 768
 FT DISULFID 773 814
 FT DISULFID 800 827
 FT DISULFID 832 877
 FT DISULFID 863 890
 FT DISULFID 894 939
 FT DISULFID 922 952
 FT CARBOHYD 376 376
 FT CARBOHYD 525 525
 FT CARBOHYD 605 605
 FT CARBOHYD 620 620
 FT CARBOHYD 752 752
 FT CARBOHYD 789 789
 FT VARSPIC 529 553
 FT VARSPIC 892 958
 SQ SEQUENCE 958 AA; 107027 MW; 416125885ABC764 CRC64;
 Query Match 8.3%; Score 87; DB 1; Length 958;
 Best Local Similarity 19.7%; Pred. No. 2.3;
 Matches 47; Conservative 24; Mismatches 82; Indels 86; Gaps 8;
 QY 6 FKNDATETIPELGEYEPPEPELENN----- 30
 DB 343 FKGDLGQLPPESGIGPEPEPLADQLQYGNQSSARVALLMQRKRSRTAGALS 402
 QY 31 -----KTMNRAENG-----GRPHHPETK-DVSEYSCRELFHFRVYTDGPCRS 73
 DB 403 RKGGGDDSSKTSKXKGIYDEEAGYTPHPDPEPEPEPEPEVYDIQGFTE----- 457
 QY 74 AKVTELVCSGCCGP-----ARLLPNAIGRKMWRPSPDPFRCPDPRYRAORVQLCPGGEA 130
 DB 458 -----VSEIRPFGIIGPMGBRCLCKIRCVAKWVG-----LCATNEE 495
 QY 131 PRARKYR--LVASCKCRRLRFRNQSLEKDFGTBAPPOKGRPRPARAKANQVEL 186
 DB 496 DDNGVVKFQPLYSCHVNRIPS-HLLSYRNISVTPPIPPNMGWKRTRLSKSTLLSTNEI 553
 RESULT 10
 ID DAN_HUMAN STANDARD; PRT; 180 AA.
 AC P41271;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Neuroblastoma suppressor of tumorigenicity 1 (Zinc finger protein
 DE DAN) (N03).
 GN NBL1 OR DAN.

OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 RN NCBI_TaxID=9606;
 RP [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lung;
 RX MEDLINE=9436724; PubMed=8084583;
 RA Enomoto H., Ozaki T., Takahashi E., Nomura N., Tabata S.,
 RA Takahashi H., Ohnuma N., Tanabe M., Iwai J., Yoshida H., Matsunaga T.,
 RA Sakiyama S.;
 RT "Identification of human DAN gene, mapping to the putative
 RT neuroblastoma tumor suppressor locus.";
 RL Oncogene 9:2785-2791(1994).
 CC -1- FUNCTION: POSSIBLE CANDIDATE AS A TUMOR SUPPRESSOR GENE OF
 CC NEUROBLASTOMA. MAY PLAY AN IMPORTANT ROLE IN PREVENTING CELLS
 CC FROM ENTERING THE FINAL STAGE (G1/S) OF THE TRANSFORMATION
 CC PROCESS.
 CC -1- TISSUE SPECIFICITY: MOST ABUNDANT IN NORMAL LUNG AND MENINGIOMA.
 CC -1- DISEASE: THE LOSS OF DAN GENE BY MUTATION IS POSSIBLY THE CAUSE OF
 CC THE DEVELOPMENT AND/OR PROGRESSION OF HUMAN NEUROBLASTOMA.
 CC -1- SIMILARITY: HIGH, TO OTHER MAMMALIAN DAN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; D28124; BAA05671.1; -;
 DR Genew; HGNC:7650; NBL1.
 DR MIM: 600613; -;
 DR InterPro: IPR000359; Cys knot.
 DR InterPro: IPR004133; DAN_dom.
 DR Pfam; PF03045; DAN; 1.
 DR SMART; SM00041; CT; 1.
 DR Anti-oncogene; DNA-binding; Zinc-finger.
 DR ZN FING 81 101
 FT DOMAIN 141 171 PRO-RICH (POTENTIAL).
 FT SEQUENCE 180 AA; 19277 MW; 15233229DB65865 CRC64;
 SQ
 Query Match 8.1%; Score 85; DB 1; Length 180;
 Best Local Similarity 26.0%; Pred. No. 0.63;
 Matches 33; Conservative 15; Mismatches 61; Indels 18; Gaps 7;
 QY 57 CRELHTRVYTDGPGCRSAKPVTELVCSCGCGPARLLPNAIGRGKMRPSPGPDF-----RC 111
 DB 34 CEAKNITQIVGSHGC-EAKSIQNRACIGQCF-SYVYNTF-----POSTSLVHCDSG 84
 QY 112 IPRVYAGRVOLLCPG-GEAPRARK-VRLVASCKCKRLTRFNQSELKXPF--GTEARQ 167
 DB 85 MPASQWEIVTLECFHBEVPRVDKLVKILKSCQACGKPSHEGLSYVQGEDGPGSQ 144
 QY 168 KGRKRP 174
 DB 145 PGTNPR 151
 RESULT 11
 LSHB TRIUV STANDARD; PRT; 141 AA.
 ID LSHB TRIUV
 AC 046482;
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DE Lutropin beta chain precursor [luteinizing hormone beta subunit] (LSH-
 DE beta) (LSH-B) (LH-B).
 GN LHB.
 OS Trichosurus vulpecula (Brush-tailed possum).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Metatheria; Diprotodontia; Phalangeridae; Trichosurus.

OX NCBI_TaxID=9337;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Plutary;
 RX MEDLINE=98345424; PubMed=9680384;
 RA Harrison G.A., Deane E.M., Cooper D.W.;
 RT "cDNA cloning of luteinizing hormone subunits from brushtail possum
 RT and red kangaroo.";
 RL Mamm. Genome 9:638-642(1998).
 RP [2]
 RP SEQUENCE FROM N.A.
 RA Lawrence S.B., McNally K.P., Fidler A.E.;
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: PROMOTES SPERMATOGENESIS AND OVULATION BY STIMULATING
 CC THE TESTES AND OVARIES TO SYNTHESIZE STEROIDS.
 CC -1- SUBUNIT: HETEROIMER OF A COMMON ALPHA CHAIN AND A UNIQUE BETA
 CC CHAIN WHICH CONFORMS BIOLOGICAL SPECIFICITY TO THYROTROPIN,
 CC LUTROPIN, FOLLITROPIN AND GONADOTROPIN.
 CC -1- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES BETA CHAIN
 CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF017448; AAC96019.1; -;
 DR EMBL; AF090388; AAC63526.1; -;
 DR HSSP; P01233; IXUL.
 DR InterPro: IPR000359; Cys knot.
 DR InterPro: IPR001545; Gly hormone.
 DR Pfam; PF00007; Cys knot; 1.
 DR SMART; SM00068; GHE; 1.
 DR PROSITE; PS00261; GLYCO_HORMONE_BETA_1; 1.
 DR PROSITE; PS00689; GLYCO_HORMONE_BETA_2; 1.
 DR Hormone; signal; Glycoprotein.
 FT SIGNAL 1 22
 FT CHAIN 23 141 LUTROPIN BETA CHAIN.
 FT DISULFID 30 78 BY SIMILARITY.
 FT DISULFID 44 93 BY SIMILARITY.
 FT DISULFID 47 131 BY SIMILARITY.
 FT DISULFID 55 109 BY SIMILARITY.
 FT DISULFID 59 111 BY SIMILARITY.
 FT DISULFID 114 121 BY SIMILARITY.
 FT CARBOHYD 34 34 N-LINKED (GLCNAC...) (POTENTIAL).
 FT SEQUENCE 141 AA; 15060 MW; C6CF98036B3C4EB0 CRC64;
 SQ
 Query Match 8.0%; Score 83.5; DB 1; Length 141;
 Best Local Similarity 27.7%; Pred. No. 0.67;
 Matches 33; Conservative 10; Mismatches 59; Indels 17; Gaps 5;
 QY 57 CRELHTRVYTDGPGCRSAKPVTELVCSCGCGPARLLPNAIGRGKMRPSPGPDFRCIPDR 115
 DB 30 CRPTVATLAESDACPVCVTFITICAGYCPSPVRLPAL-----PSPQVLCVYRE 82
 QY 116 YRAQVOLL-CPGGEAP-RARKVRLVASCKCKRLTRFNQSELKDFGTGARPPKGRK 172
 DB 83 LSFSSIRLPGCPGVDPITFSFVALSCGSCGRSLH-----SDCGPRPRLCTRP 134
 RESULT 12
 BCAL HUMAN STANDARD; PRT; 870 AA.
 ID BCAL HUMAN
 AC P56945;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE CRK-associated substrate (p130Cas) (Breast cancer anti-estrogen
 DE resistance 1 protein).
 GN BCAR1 OR CRKAS OR CAS.

OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 CC NCBI_TaxID=9606;
 RN
 RP SEQUENCE FROM N.A.
 RC TISSUE=Breast cancer;
 RX MEDLINE=20119429; PubMed=10639512;
 RA Brinkman A., van der Flier S., Kok E.M., Dorssers L.C.J.;
 RT "BCAR1, a human homologue of the adapter protein p130Cas, induces
 RT anti-estrogen resistance in breast cancer cells.";
 RL J. Natl. Cancer Inst. 92:112-120(2000).
 RN
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RA Otto E., Birnbaum S., Verbeek M., Hildebrandt F.;
 RT "Interaction between human Crk-associated substrate (p130Cas) and
 RT nephrin.";
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 RN
 RP SEQUENCE FROM N.A.
 RC TISSUE=Cornea;
 RA Imoto Y., Ohguro N., Yoshida A., Tsujikawa M., Inoue Y., Tano Y.;
 RT "The effects of growth factors on tyrosine phosphorylation of p130Cas
 RT in corneal epithelial cell.";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
 RA Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshikawa Y.,
 RA Matsunawa H., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A.,
 RA Nakamura Y., Nagahara K., Masuko Y., Sasaki N.;
 RT "NEDD human cDNA sequencing project.";
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 CC
 CC -1- FUNCTION: DOCKING PROTEIN WHICH PLAYS A CENTRAL COORDINATING ROLE
 CC FOR TYROSINE-KINASE-BASED SIGNALING RELATED TO CELL ADHESION.
 CC IMPLICATED IN INDUCTION OF CELL MIGRATION, OVEREXPRESSION CONFERS
 CC ANTITUMOR RESISTANCE ON BREAST CANCER CELLS.
 CC
 CC -1- SUBUNIT: FORMS COMPLEXES IN VIVO WITH FOCAL ADHESION KINASE 1,
 CC ADAPTER PROTEIN CRKL AND LYN KINASE. CAN HETERODIMERIZE WITH CASL
 CC INTERACTS WITH NEPHROCISTIN AND PTK2B (BY SIMILARITY).
 CC
 CC -1- SUBCELLULAR LOCATION: FOCAL ADHESIONS. UNPHOSPHORYLATED FORM
 CC LOCALIZES IN THE CYTOPLASM AND CAN MOVE TO THE MEMBRANE UPON
 CC TYROSINE PHOSPHORYLATION (BY SIMILARITY).
 CC
 CC -1- TISSUE SPECIFICITY: WIDELY EXPRESSED WITH AN ABUNDANT EXPRESSION
 CC IN THE TESTIS. LOW LEVEL OF EXPRESSION SEEN IN THE LIVER, THYMUS,
 CC AND PERIPHERAL BLOOD LEUCOCYTES. THE PROTEIN HAS BEEN DETECTED IN
 CC A B-CELL LINE.
 CC
 CC -1- DOMAIN: CONTAINS A CENTRAL DOMAIN (SUBSTRATE DOMAIN) CONTAINING
 CC MULTIPLE POTENTIAL SH2-BINDING SITES AND A C-TERMINAL DOMAIN
 CC CONTAINING A DIVERGENT HELIX-LOOP-HELIX (HLH) MOTIF. THE SH2-
 CC BINDING SITES PUTATIVELY BIND CRK, NCK AND ABL SH2 DOMAINS.
 CC THE HLH MOTIF IS ABSOLUTELY REQUIRED FOR THE INDUCTION OF
 CC PSEUDOPODIAL GROWTH IN YEAST AND MEDIATES HETERODIMERIZATION
 CC WITH CASL (BY SIMILARITY).
 CC
 CC -1- DOMAIN: A SERINE-RICH REGION PROMOTES ACTIVATION OF THE SERUM
 CC RESPONSE ELEMENT (SRE).
 CC
 CC -1- DOMAIN: THE SH3 DOMAIN IS NECESSARY FOR THE LOCALIZATION OF THE
 CC PROTEIN TO FOCAL ADHESIONS AND INTERACTS WITH ONE PROLINE-RICH
 CC REGION OF FOCAL ADHESION KINASE 1.
 CC
 CC -1- PTM: FOCAL ADHESION KINASE 1 PHOSPHORYLATES THE PROTEIN AT THE
 CC YDYVL MOTIF. THE SRC FAMILY KINASES ARE RECRUITED TO THE
 CC PHOSPHORYLATED SITES AND CAN PHOSPHORYLATE OTHER TYROSINE
 CC RESIDUES. TYROSINE PHOSPHORYLATION IS TRIGGERED BY INTEGRIN
 CC MEDIATED ADHESION OF CELLS TO THE EXTRACELLULAR MATRIX.
 CC
 CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
 CC
 CC -1- SIMILARITY: BELONGS TO THE CAS FAMILY.
 CC
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC
 CC DR EMBL; AJ242987; CAB75875.2; -
 CC DR EMBL; AF218451; AAF27527.1; -
 CC DR EMBL; AB040024; BAA92711.1; -
 CC DR EMBL; AK027608; BAB55230.1; -
 CC DR HSSP; P07751; 1BK2.
 CC DR Genew; HGNC:971; BCAR1.
 CC DR MIM; 602941; -
 CC DR InterPro; IPR001452; SH3.
 CC DR Pfam; PF00018; SH3; 1.
 CC DR PRINTS; PR00452; SH3DOMAIN.
 CC DR ProDom; PD000066; SH3; 1.
 CC DR SMART; SM00326; SH3; 1.
 CC DR PROSITE; PS50002; SH3; 1.
 CC KW Phosphorylation; SH3 domain; SH3-binding; Cell adhesion.
 CC FT DOMAIN 3 65
 CC FT DOMAIN 74 87
 CC FT DOMAIN 115 416
 CC FT DOMAIN 422 614
 CC FT SITE 635 643
 CC FT DOMAIN 746 796
 CC FT CONFLICT 1 4
 CC FT CONFLICT 63 63
 CC FT CONFLICT 236 236
 CC FT CONFLICT 349 349
 CC FT CONFLICT 363 363
 CC FT CONFLICT 428 428
 CC FT CONFLICT 471 471
 CC SQ SEQUENCE 870 AA; 93361 MW; D57A9CAFCA4FAD0 CRC64;
 CC
 CC Query Match 7.8%; Score 81.5; DB 1; Length 870;
 CC Best Local Similarity 20.9%; Pred. No. 6.7;
 CC Matches 41; Conservative 27; Mismatches 71; Indels 57; Gaps 9;
 CC
 CC QY 9 DATEIIP-----LGYPEPPPELENNKTMNAENGRRPHHPFTKDVSESCRELH 61
 CC DB 112 DSVIVLPSPKAQGLYVPGSPSPQSPAKQSTSTSKQTHHPFSPADLYQ----- 166
 CC QY 62 FTRYVTDGCRSAKVELTVCSGQG-----PAR-LLPNAIGRK 100
 CC DB 167 -----VPPGGGPAQDIYVPPSAGSHDIYVPPSMDSRWEGTKPRAKVVPFRVGGY 222
 CC QY 101 WWRPSGPPRCPLPRYKQVQLCPGG-----EAPRAKALVASCKKLTTRHNGSEL 156
 CC DB 223 VYEAQPE-----QDEYDIPR-HLAPQDIDYVPPVPGT-----LPQYGO-EV 266
 CC QY 157 KDFGTEAARPKGRKP 172
 CC DB 267 YTPPMAYKGPNGRDP 282
 CC
 CC RESULT 13
 CC GTSR HUMAN
 CC ID GTSR HUMAN STANDARD; PRT; 720 AA.
 CC AC Q9NTZ3; Q9BRE0; Q9Y57; Q9UGZ9;
 CC DT 15-JUN-2002 (Rel. 41, Created)
 CC DT 15-JUN-2002 (Rel. 41, Last sequence update)
 CC DT 15-JUN-2002 (Rel. 41, Last annotation update)
 CC DE G2 and S phase expressed protein 1 (B99 homolog).
 CC GN GTSR1.
 CC OS Homo sapiens (Human).
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 CC OK NCBI_TaxID=9606;
 CC RN
 CC RP SEQUENCE FROM N.A.
 CC RC TISSUE=Placenta;
 CC RX MEDLINE=20432105; PubMed=10974554;
 CC RA Monte M., Collavin L., Lazarevic D., Utrera R., Dragani T.A.,
 CC Schneider C.;

"Cloning, chromosome mapping and functional characterization of a human homologue of murine Gtse-1 (B99) gene.";
 RT Gene 254:229-236(2000).
 RN [2]
 RX SEQUENCE FROM N.A. PubMed=10591208;
 RA MEDLINE=20057165; PubMed=10591208;
 RA Dunham I., Hunt A.R., Collins J.E., Almeida J.P., Babapae A.K., Clump M., Smith L.J., Ainscough R., Almeida J.P., Babapae A.K., Bagley C., Bailey J., Barlow K.F., Bates K.N., Beasley O.P., Bird C.P., Blakey S.E., Bridgman A.M., Buck D., Burgess J., Burrill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G., Clegg S.M., Cobley V.E., Cole C.G., Collier R.E., Connor R., Conroy D., Corby N.R., Coville G.J., Cox A.V., Davis J., Dawson E., Dhan P.D., Dockree C., Dodsworth S.D., Durbin R.M., Ellington A.G., Evans K.L., Fey J.M., Fleming K., French L., Garner A.A., Gilbert J.G.R., Goward M.E., Graham D.V., Griffiths M.N.D., Hall C., Hall R.E., Hall-Tamlyn G., Heathcote R.W., Ho S., Holmes S., Hunt S.E., Jones M.C., Kershaw J., Kimberley A.M., King A., Laid G.K., Langford C.F., Leverhwa M.A., Lloyd C., Lloyd D.M., Maryn I.D., Mashreghi-Mohammadi M., Matthews L.H., Mccann O.T., Mcclary J., McLaren S., McMuray A.A., Milne S.A., Morrison B.J., Odell C.N., Pavitt R., Pearce A.V., Pearson S., Phillimore B.J.C.T., Phillips S.H., Plumb R.W., Ramsay H., Ramsey Y., Rogers L., Ross M.T., Scott C.E., Sehra H.K., Skuce C.D., Smalley S., Smith M.L., Soderlund C., Spragon L., Steward C.A., Sulston J.E., Swann R.M., Vaudin M., Wall M., Wallis J.M., Whiteley M.N., Willey D.L., Williams L., Williams S.A., Williamson H., Wilmer T.E., Wilming L., Wright C.L., Hubbard T., Bentley D.R., Beck S., Rogers J., Shimizu N., Minoshima S., Kawasaki K., Sasaki T., Asakawa S., Kudoh J., Shintani A., Shibuya K., Yoshizaki Y., Aoki N., Mitsuama S., Roe B.A., Chen F., Chu Y., Hu P., Hua A., Kenton S., Lai H., Lao H.I., Dorman A., Fang F., Fu Y., Hu P., Hua A., Kenton S., Lai H., Lao H.I., Lewis J., Lewis S., Lin S.-P., Loh P., Malai E., Nguyen T., Pan H., Phan S., Qi S., Qian Y., Ray L., Ren Q., Shaull S., Sloan D., Song L., Zhang Q., Wang Y., Wang Z., White J., Willingham D., Wu H., Yao Z., Zhan M., Zhang G., Chisoe S., Murray J., Miller N., Min P., Fulan M., Johnson D., Bemis G., Bentley D., Bradsenaw H., Bourne S., Cordes M., Du Z., Fulton L., Goela D., Graves T., Hawkins J., Hinds K., Kemp K., Latreille P., Layman D., Ozeresky P., Rohling T., Scheet P., Walker C., Wamsley A., Woldmann P., Pepin K., Nelson J., Korn I., Bedell J.A., Hillier L., Mardis E., Waterston R., Wilson R., Emanuel B.S., Shaikh T., Kurahashi H., Salter S., Budarf M.L., McDermid H.E., Johnson A., Wong A.C.C., Morrow B.E., Edelmann L., Kim U.-J., Shinya H., Simon M.I., Dumanski J.P., Payard M., Kedra D., Serousi E., Franssen I., Tapia I., Bruder C.E., O'Brien K.P., Wilkinson P., Bodenteich A., Hartman K., Hu X., Khan A.S., Lane L., Tilahun Y., Wright H.;
 RT "The DNA sequence of human chromosome 22.";
 RT Nature 402:489-495(1999).
 RN [3]
 RX SEQUENCE FROM N.A.
 RA TISSUE=Muscle;
 RA Strausberg R.;
 RA Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: May be involved in p53-induced cell cycle arrest in G2/M phase by interfering with microtubule rearrangements that are required to enter mitosis. Overexpression delays G2/M phase progression.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic. Associated with microtubules.
 CC -1- DEVELOPMENTAL STAGE: Expressed in G2/M phase. Not detected in quiescent cells.
 CC -1- PTM: Phosphorylated in mitosis (By similarity).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC EMBL, AF223408; AAF31459.1; -;
 DR EMBL, AL031588; CAB38415.1; -;

DR EMBL, AL022325; CAB63079.1; -;
 DR EMBL, BC006325; AAH06325.1; -;
 KM Microtubules; Phosphorylation.
 FT DOMAIN 22 27 POLY-SER.
 FT CONFLICT 259 259 V -> I (IN REF. 1).
 FT CONFLICT 506 506 R -> W (IN REF. 2; CAB38415).
 SQ SEQUENCE 720 AA; 76614 MW; ACD91CCDD008A89C CRC64;
 Query March 7.6%; Score 80; DB 1; Length 720;
 Best Local Similarity 21.1%; Pred. No. 7.5;
 Matches 32; Conservative 21; Mismatches 39; Indels 60; Gaps 6;
 QY 70 PCRSAPVTELVCSGQCSPARLP-----NAIGRGKWRPSPGDFRCIPDRYARV----- 121
 DB 341 PANSRPLSNISKSGRMGPALRLPALRPGVPGASSW-----QAKRVDS 385
 QY 122 ---QLCPGSEAP-----PARKRLVASC---KCKRLT 148
 DB 386 LAEQLTPAPASFPQPTPGCGQWINSQAWSESSQLNTRIRRRSDCLNSKTYMP 445
 QY 149 RPHNSELKDF-----GTBARPKGRKRPRA 176
 DB 446 TPTNQFKLPKFSIGDSPSSTPKLSRAQRPQS 477
 RESULT 14
 SLIT DROME STANDARD; PRT: 1480 AA.
 ID SLIT DROME
 AC P24014;
 DT 01-MAR-1992 (Rel. 21, Created)
 DT 01-MAR-1992 (Rel. 21, Last sequence update)
 DT 15-JUN-2002 (Rel. 21, Last annotation update)
 DE Slit protein precursor.
 GN SLIT
 OS Drosophila melanogaster (fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 CC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 CC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 CX NCBI_TaxID=7227;
 RN RN
 RX SEQUENCE FROM N.A.
 RA MEDLINE=9109665; PubMed=2176636;
 RA Rothenberg J.M., Jacobs J.R., Goodman C.S., Artavanis-Tsakonas S.;
 RT "Slit: an extracellular protein necessary for development of midline glia and commissural axon pathways contains both EGF and LRR domains.";
 RT Genes Dev. 4:2169-2187(1990).
 RL Genes Dev. 4:2169-2187(1990).
 CC -1- FUNCTION: NECESSARY FOR DEVELOPMENT OF MIDLINE GLIA AND COMMISSURAL AXON PATHWAYS. SLIT MAY INTERACT WITH EXTRACELLULAR MATRIX MOLECULES.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: A LONG FORM (SHOWN HERE) AND A SHORT FORM, ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: EXCRETED BY THE MIDLINE GLIA CELLS AND EVENTUALLY DISTRIBUTED ALONG THE AXONS.
 CC -1- SIMILARITY: CONTAINS 7 EGF-LIKE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 24 LEUCINE-RICH REPEATS (LRR).
 CC -1- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTEINE KNOT-LIKE (CTCK) DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 LAMININ G-LIKE DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC EMBL, X53959; CAA37910.1; -;
 DR HSSP, P00740; 1EDM.
 DR PIR, A36665; A36665.
 DR FLYBase, FBgn003425; slt.
 DR InterPro, IPR000152; Aex_hydroxyl.
 DR InterPro, IPR000359; Cys_knot.

DR InterPro; IPR000561; EGF-like.
 DR InterPro; IPR000742; EGF_2.
 DR InterPro; IPR001881; EGF_Ca.
 DR InterPro; IPR001611; LRR_Nterm.
 DR InterPro; IPR000483; LRR_Cterm.
 DR InterPro; IPR000372; LRR_Nterm.
 DR InterPro; IPR003592; LRR_out.
 DR InterPro; IPR003591; LRR_Typ.
 DR InterPro; IPR001791; Laminin_G.
 DR Pfam; PF000007; Cys_knot; 1.
 DR Pfam; PF000008; EGF_7.
 DR Pfam; PF00054; Laminin_G; 1.
 DR Pfam; PF00560; LRR; 16.
 DR Pfam; PF01462; LRRNT; 4.
 DR Pfam; PF01463; LRRCT; 4.
 DR SMART; SM00041; CT; 1.
 DR SMART; SM00179; EGF_Ca; 2.
 DR SMART; SM00001; EGF_Like; 5.
 DR SMART; SM00370; LRR; 4.
 DR SMART; SM00082; LRRCT; 4.
 DR SMART; SM00013; LRRNT; 4.
 DR SMART; SM00369; LRR_Typ; 9.
 DR SMART; SM00282; LamG; 1.
 DR PROSITE; PS00010; ASX_HYDROXYL; 3.
 DR PROSITE; PS01185; CTCK_1; 1.
 DR PROSITE; PS01225; CTCK_2; 1.
 DR PROSITE; PS00022; EGF_1; 7.
 DR PROSITE; PS01186; EGF_2; 5.
 DR PROSITE; PS01187; EGF_Ca; 2.
 DR PROSITE; PS50025; LAM_G_DOMAIN; 1.
 KW Neurogenesis1; Glycoprotein; Signal; Alternative splicing;
 KW EGF-like domain; Repeat; Leucine-rich repeat.
 FT SIGNAL 1 36
 FT CHAIN 37 1480 SLIT PROTEIN.
 FT REPEAT 99 122 LRR 1.
 FT REPEAT 123 146 LRR 2.
 FT REPEAT 148 170 LRR 3.
 FT REPEAT 171 194 LRR 4.
 FT REPEAT 195 218 LRR 5.
 FT REPEAT 220 246 LRR 6.
 FT REPEAT 321 344 LRR 7.
 FT REPEAT 345 368 LRR 8.
 FT REPEAT 369 392 LRR 9.
 FT REPEAT 394 416 LRR 10.
 FT REPEAT 417 440 LRR 11.
 FT REPEAT 422 445 LRR 12.
 FT REPEAT 452 475 LRR 13.
 FT REPEAT 470 493 LRR 14.
 FT REPEAT 500 523 LRR 15.
 FT REPEAT 523 546 LRR 16.
 FT REPEAT 546 569 LRR 17.
 FT REPEAT 569 593 LRR 18.
 FT REPEAT 593 617 LRR 19.
 FT REPEAT 617 641 LRR 20.
 FT REPEAT 641 666 LRR 21.
 FT REPEAT 666 690 LRR 22.
 FT REPEAT 690 714 LRR 23.
 FT REPEAT 714 738 LRR 24.
 FT REPEAT 738 762 LRR 25.
 FT REPEAT 762 786 LRR 26.
 FT REPEAT 786 810 LRR 27.
 FT REPEAT 810 834 LRR 28.
 FT REPEAT 834 858 LRR 29.
 FT REPEAT 858 882 LRR 30.
 FT REPEAT 882 906 LRR 31.
 FT REPEAT 906 930 LRR 32.
 FT REPEAT 930 954 LRR 33.
 FT REPEAT 954 978 LRR 34.
 FT REPEAT 978 1002 LRR 35.
 FT REPEAT 1002 1026 LRR 36.
 FT REPEAT 1026 1050 LRR 37.
 FT REPEAT 1050 1074 LRR 38.
 FT REPEAT 1074 1098 LRR 39.
 FT REPEAT 1098 1122 LRR 40.
 FT REPEAT 1122 1146 LRR 41.
 FT REPEAT 1146 1170 LRR 42.
 FT REPEAT 1170 1194 LRR 43.
 FT REPEAT 1194 1218 LRR 44.
 FT REPEAT 1218 1242 LRR 45.
 FT REPEAT 1242 1266 LRR 46.
 FT REPEAT 1266 1290 LRR 47.
 FT REPEAT 1290 1314 LRR 48.
 FT REPEAT 1314 1338 LRR 49.
 FT REPEAT 1338 1362 LRR 50.
 FT REPEAT 1362 1386 LRR 51.
 FT REPEAT 1386 1410 LRR 52.
 FT REPEAT 1410 1434 LRR 53.
 FT REPEAT 1434 1458 LRR 54.
 FT REPEAT 1458 1482 LRR 55.
 FT REPEAT 1482 1506 LRR 56.
 FT REPEAT 1506 1530 LRR 57.
 FT REPEAT 1530 1554 LRR 58.
 FT REPEAT 1554 1578 LRR 59.
 FT REPEAT 1578 1602 LRR 60.
 FT REPEAT 1602 1626 LRR 61.
 FT REPEAT 1626 1650 LRR 62.
 FT REPEAT 1650 1674 LRR 63.
 FT REPEAT 1674 1698 LRR 64.
 FT REPEAT 1698 1722 LRR 65.
 FT REPEAT 1722 1746 LRR 66.
 FT REPEAT 1746 1770 LRR 67.
 FT REPEAT 1770 1794 LRR 68.
 FT REPEAT 1794 1818 LRR 69.
 FT REPEAT 1818 1842 LRR 70.
 FT REPEAT 1842 1866 LRR 71.
 FT REPEAT 1866 1890 LRR 72.
 FT REPEAT 1890 1914 LRR 73.
 FT REPEAT 1914 1938 LRR 74.
 FT REPEAT 1938 1962 LRR 75.
 FT REPEAT 1962 1986 LRR 76.
 FT REPEAT 1986 2010 LRR 77.
 FT REPEAT 2010 2034 LRR 78.
 FT REPEAT 2034 2058 LRR 79.
 FT REPEAT 2058 2082 LRR 80.
 FT REPEAT 2082 2106 LRR 81.
 FT REPEAT 2106 2130 LRR 82.
 FT REPEAT 2130 2154 LRR 83.
 FT REPEAT 2154 2178 LRR 84.
 FT REPEAT 2178 2202 LRR 85.
 FT REPEAT 2202 2226 LRR 86.
 FT REPEAT 2226 2250 LRR 87.
 FT REPEAT 2250 2274 LRR 88.
 FT REPEAT 2274 2298 LRR 89.
 FT REPEAT 2298 2322 LRR 90.
 FT REPEAT 2322 2346 LRR 91.
 FT REPEAT 2346 2370 LRR 92.
 FT REPEAT 2370 2394 LRR 93.
 FT REPEAT 2394 2418 LRR 94.
 FT REPEAT 2418 2442 LRR 95.
 FT REPEAT 2442 2466 LRR 96.
 FT REPEAT 2466 2490 LRR 97.
 FT REPEAT 2490 2514 LRR 98.
 FT REPEAT 2514 2538 LRR 99.
 FT REPEAT 2538 2562 LRR 100.
 FT REPEAT 2562 2586 LRR 101.
 FT REPEAT 2586 2610 LRR 102.
 FT REPEAT 2610 2634 LRR 103.
 FT REPEAT 2634 2658 LRR 104.
 FT REPEAT 2658 2682 LRR 105.
 FT REPEAT 2682 2706 LRR 106.
 FT REPEAT 2706 2730 LRR 107.
 FT REPEAT 2730 2754 LRR 108.
 FT REPEAT 2754 2778 LRR 109.
 FT REPEAT 2778 2802 LRR 110.
 FT REPEAT 2802 2826 LRR 111.
 FT REPEAT 2826 2850 LRR 112.
 FT REPEAT 2850 2874 LRR 113.
 FT REPEAT 2874 2898 LRR 114.
 FT REPEAT 2898 2922 LRR 115.
 FT REPEAT 2922 2946 LRR 116.
 FT REPEAT 2946 2970 LRR 117.
 FT REPEAT 2970 2994 LRR 118.
 FT REPEAT 2994 3018 LRR 119.
 FT REPEAT 3018 3042 LRR 120.
 FT REPEAT 3042 3066 LRR 121.
 FT REPEAT 3066 3090 LRR 122.
 FT REPEAT 3090 3114 LRR 123.
 FT REPEAT 3114 3138 LRR 124.
 FT REPEAT 3138 3162 LRR 125.
 FT REPEAT 3162 3186 LRR 126.
 FT REPEAT 3186 3210 LRR 127.
 FT REPEAT 3210 3234 LRR 128.
 FT REPEAT 3234 3258 LRR 129.
 FT REPEAT 3258 3282 LRR 130.
 FT REPEAT 3282 3306 LRR 131.
 FT REPEAT 3306 3330 LRR 132.
 FT REPEAT 3330 3354 LRR 133.
 FT REPEAT 3354 3378 LRR 134.
 FT REPEAT 3378 3402 LRR 135.
 FT REPEAT 3402 3426 LRR 136.
 FT REPEAT 3426 3450 LRR 137.
 FT REPEAT 3450 3474 LRR 138.
 FT REPEAT 3474 3498 LRR 139.
 FT REPEAT 3498 3522 LRR 140.
 FT REPEAT 3522 3546 LRR 141.
 FT REPEAT 3546 3570 LRR 142.
 FT REPEAT 3570 3594 LRR 143.
 FT REPEAT 3594 3618 LRR 144.
 FT REPEAT 3618 3642 LRR 145.
 FT REPEAT 3642 3666 LRR 146.
 FT REPEAT 3666 3690 LRR 147.
 FT REPEAT 3690 3714 LRR 148.
 FT REPEAT 3714 3738 LRR 149.
 FT REPEAT 3738 3762 LRR 150.
 FT REPEAT 3762 3786 LRR 151.
 FT REPEAT 3786 3810 LRR 152.
 FT REPEAT 3810 3834 LRR 153.
 FT REPEAT 3834 3858 LRR 154.
 FT REPEAT 3858 3882 LRR 155.
 FT REPEAT 3882 3906 LRR 156.
 FT REPEAT 3906 3930 LRR 157.
 FT REPEAT 3930 3954 LRR 158.
 FT REPEAT 3954 3978 LRR 159.
 FT REPEAT 3978 4002 LRR 160.
 FT REPEAT 4002 4026 LRR 161.
 FT REPEAT 4026 4050 LRR 162.
 FT REPEAT 4050 4074 LRR 163.
 FT REPEAT 4074 4098 LRR 164.
 FT REPEAT 4098 4122 LRR 165.
 FT REPEAT 4122 4146 LRR 166.
 FT REPEAT 4146 4170 LRR 167.
 FT REPEAT 4170 4194 LRR 168.
 FT REPEAT 4194 4218 LRR 169.
 FT REPEAT 4218 4242 LRR 170.
 FT REPEAT 4242 4266 LRR 171.
 FT REPEAT 4266 4290 LRR 172.
 FT REPEAT 4290 4314 LRR 173.
 FT REPEAT 4314 4338 LRR 174.
 FT REPEAT 4338 4362 LRR 175.
 FT REPEAT 4362 4386 LRR 176.
 FT REPEAT 4386 4410 LRR 177.
 FT REPEAT 4410 4434 LRR 178.
 FT REPEAT 4434 4458 LRR 179.
 FT REPEAT 4458 4482 LRR 180.
 FT REPEAT 4482 4506 LRR 181.
 FT REPEAT 4506 4530 LRR 182.
 FT REPEAT 4530 4554 LRR 183.
 FT REPEAT 4554 4578 LRR 184.
 FT REPEAT 4578 4602 LRR 185.
 FT REPEAT 4602 4626 LRR 186.
 FT REPEAT 4626 4650 LRR 187.
 FT REPEAT 4650 4674 LRR 188.
 FT REPEAT 4674 4698 LRR 189.
 FT REPEAT 4698 4722 LRR 190.
 FT REPEAT 4722 4746 LRR 191.
 FT REPEAT 4746 4770 LRR 192.
 FT REPEAT 4770 4794 LRR 193.
 FT REPEAT 4794 4818 LRR 194.
 FT REPEAT 4818 4842 LRR 195.
 FT REPEAT 4842 4866 LRR 196.
 FT REPEAT 4866 4890 LRR 197.
 FT REPEAT 4890 4914 LRR 198.
 FT REPEAT 4914 4938 LRR 199.
 FT REPEAT 4938 4962 LRR 200.
 FT REPEAT 4962 4986 LRR 201.
 FT REPEAT 4986 5010 LRR 202.
 FT REPEAT 5010 5034 LRR 203.
 FT REPEAT 5034 5058 LRR 204.
 FT REPEAT 5058 5082 LRR 205.
 FT REPEAT 5082 5106 LRR 206.
 FT REPEAT 5106 5130 LRR 207.
 FT REPEAT 5130 5154 LRR 208.
 FT REPEAT 5154 5178 LRR 209.
 FT REPEAT 5178 5202 LRR 210.
 FT REPEAT 5202 5226 LRR 211.
 FT REPEAT 5226 5250 LRR 212.
 FT REPEAT 5250 5274 LRR 213.
 FT REPEAT 5274 5298 LRR 214.
 FT REPEAT 5298 5322 LRR 215.
 FT REPEAT 5322 5346 LRR 216.
 FT REPEAT 5346 5370 LRR 217.
 FT REPEAT 5370 5394 LRR 218.
 FT REPEAT 5394 5418 LRR 219.
 FT REPEAT 5418 5442 LRR 220.
 FT REPEAT 5442 5466 LRR 221.
 FT REPEAT 5466 5490 LRR 222.
 FT REPEAT 5490 5514 LRR 223.
 FT REPEAT 5514 5538 LRR 224.
 FT REPEAT 5538 5562 LRR 225.
 FT REPEAT 5562 5586 LRR 226.
 FT REPEAT 5586 5610 LRR 227.
 FT REPEAT 5610 5634 LRR 228.
 FT REPEAT 5634 5658 LRR 229.
 FT REPEAT 5658 5682 LRR 230.
 FT REPEAT 5682 5706 LRR 231.
 FT REPEAT 5706 5730 LRR 232.
 FT REPEAT 5730 5754 LRR 233.
 FT REPEAT 5754 5778 LRR 234.
 FT REPEAT 5778 5802 LRR 235.
 FT REPEAT 5802 5826 LRR 236.
 FT REPEAT 5826 5850 LRR 237.
 FT REPEAT 5850 5874 LRR 238.
 FT REPEAT 5874 5898 LRR 239.
 FT REPEAT 5898 5922 LRR 240.
 FT REPEAT 5922 5946 LRR 241.
 FT REPEAT 5946 5970 LRR 242.
 FT REPEAT 5970 5994 LRR 243.
 FT REPEAT 5994 6018 LRR 244.
 FT REPEAT 6018 6042 LRR 245.
 FT REPEAT 6042 6066 LRR 246.
 FT REPEAT 6066 6090 LRR 247.
 FT REPEAT 6090 6114 LRR 248.
 FT REPEAT 6114 6138 LRR 249.
 FT REPEAT 6138 6162 LRR 250.
 FT REPEAT 6162 6186 LRR 251.
 FT REPEAT 6186 6210 LRR 252.
 FT REPEAT 6210 6234 LRR 253.
 FT REPEAT 6234 6258 LRR 254.
 FT REPEAT 6258 6282 LRR 255.
 FT REPEAT 6282 6306 LRR 256.
 FT REPEAT 6306 6330 LRR 257.
 FT REPEAT 6330 6354 LRR 258.
 FT REPEAT 6354 6378 LRR 259.
 FT REPEAT 6378 6402 LRR 260.
 FT REPEAT 6402 6426 LRR 261.
 FT REPEAT 6426 6450 LRR 262.
 FT REPEAT 6450 6474 LRR 263.
 FT REPEAT 6474 6498 LRR 264.
 FT REPEAT 6498 6522 LRR 265.
 FT REPEAT 6522 6546 LRR 266.
 FT REPEAT 6546 6570 LRR 267.
 FT REPEAT 6570 6594 LRR 268.
 FT REPEAT 6594 6618 LRR 269.
 FT REPEAT 6618 6642 LRR 270.
 FT REPEAT 6642 6666 LRR 271.
 FT REPEAT 6666 6690 LRR 272.
 FT REPEAT 6690 6714 LRR 273.
 FT REPEAT 6714 6738 LRR 274.
 FT REPEAT 6738 6762 LRR 275.
 FT REPEAT 6762 6786 LRR 276.
 FT REPEAT 6786 6810 LRR 277.
 FT REPEAT 6810 6834 LRR 278.
 FT REPEAT 6834 6858 LRR 279.
 FT REPEAT 6858 6882 LRR 280.
 FT REPEAT 6882 6906 LRR 281.
 FT REPEAT 6906 6930 LRR 282.
 FT REPEAT 6930 6954 LRR 283.
 FT REPEAT 6954 6978 LRR 284.
 FT REPEAT 6978 7002 LRR 285.
 FT REPEAT 7002 7026 LRR 286.
 FT REPEAT 7026 7050 LRR 287.
 FT REPEAT 7050 7074 LRR 288.
 FT REPEAT 7074 7098 LRR 289.
 FT REPEAT 7098 7122 LRR 290.
 FT REPEAT 7122 7146 LRR 291.
 FT REPEAT 7146 7170 LRR 292.
 FT REPEAT 7170 7194 LRR 293.
 FT REPEAT 7194 7218 LRR 294.
 FT REPEAT 7218 7242 LRR 295.
 FT REPEAT 7242 7266 LRR 296.
 FT REPEAT 7266 7290 LRR 297.
 FT REPEAT 7290 7314 LRR 298.
 FT REPEAT 7314 7338 LRR 299.
 FT REPEAT 7338 7362 LRR 300.
 FT REPEAT 7362 7386 LRR 301.
 FT REPEAT 7386 7410 LRR 302.
 FT REPEAT 7410 7434 LRR 303.
 FT REPEAT 7434 7458 LRR 304.
 FT REPEAT 7458 7482 LRR 305.
 FT REPEAT 7482 7506 LRR 306.
 FT REPEAT 7506 7530 LRR 307.
 FT REPEAT 7530 7554 LRR 308.
 FT REPEAT 7554 7578 LRR 309.
 FT REPEAT 7578 7602 LRR 310.
 FT REPEAT 7602 7626 LRR 311.
 FT REPEAT 7626 7650 LRR 312.
 FT REPEAT 7650 7674 LRR 313.
 FT REPEAT 7674 7698 LRR 314.
 FT REPEAT 7698 7722 LRR 315.
 FT REPEAT 7722 7746 LRR 316.
 FT REPEAT 7746 7770 LRR 317.
 FT REPEAT 7770 7794 LRR 318.
 FT REPEAT 7794 7818 LRR 319.
 FT REPEAT 7818 7842 LRR 320.
 FT REPEAT 7842 7866 LRR 321.
 FT REPEAT 7866 7890 LRR 322.
 FT REPEAT 7890 7914 LRR 323.
 FT REPEAT 7914 7938 LRR 324.
 FT REPEAT 7938 7962 LRR 325.
 FT REPEAT 7962 7986 LRR 326.
 FT REPEAT 7986 8010 LRR 327.
 FT REPEAT 8010 8034 LRR 328.
 FT REPEAT 8034 8058 LRR 329.
 FT REPEAT 8058 8082 LRR 330.
 FT REPEAT 8082 8106 LRR 331.
 FT REPEAT 8106 8130 LRR 332.
 FT REPEAT 8130 8154 LRR 333.
 FT REPEAT 8154 8178 LRR 334.
 FT REPEAT 8178 8202 LRR 335.
 FT REPEAT 8202 8226 LRR 336.
 FT REPEAT 8226 8250 LRR 337.
 FT REPEAT 8250 8274 LRR 338.
 FT REPEAT 8274 8298 LRR 339.
 FT REPEAT 8298 8322 LRR 340.
 FT REPEAT 8322 8346 LRR 341.
 FT REPEAT 8346 8370 LRR 342.
 FT REPEAT 8370 8394 LRR 343.
 FT REPEAT 8394 8418 LRR 344.
 FT REPEAT 8418 8442 LRR 345.
 FT REPEAT 8442 8466 LRR 346.
 FT REPEAT 8466 8490 LRR 347.
 FT REPEAT 8490 8514 LRR 348.
 FT REPEAT 8514 8538 LRR 349.
 FT REPEAT 8538 8562 LRR 350.
 FT REPEAT 8562 8586 LRR 351.
 FT REPEAT 8586 8610 LRR 352.
 FT REPEAT 8610 8634 LRR 353.
 FT REPEAT 8634 8658 LRR 354.
 FT REPEAT 8658 8682 LRR 355.
 FT REPEAT 8682 8706 LRR 356.
 FT REPEAT 8706 8730 LRR 357.
 FT REPEAT 8730 8754 LRR 358.
 FT REPEAT 8754 8778 LRR 359.
 FT REPEAT 8778 8802 LRR 360.
 FT REPEAT 8802 8826 LRR 361.
 FT REPEAT 8826 8850 LRR 362.
 FT REPEAT 8850 8874 LRR 363.
 FT REPEAT 8874 8898 LRR 364.
 FT REPEAT 8898 8922 LRR 365.
 FT REPEAT 8922 8946 LRR 366.
 FT REPEAT 8946 8970 LRR 367.
 FT REPEAT 8970 8994 LRR 368.
 FT REPEAT 8994 9018 LRR 369.
 FT REPEAT 9018 9042 LRR 370.
 FT REPEAT 9042 9066 LRR 371.
 FT REPEAT 9066 9090 LRR 372.
 FT REPEAT 9090 9114 LRR 373.
 FT REPEAT 9114 9138 LRR 374.
 FT REPEAT 9138 9162 LRR 375.
 FT REPEAT 9162 9186 LRR 376.
 FT REPEAT 9186 9210 LRR 377.
 FT REPEAT 9210 9234 LRR 378.
 FT REPEAT 9234 9258 LRR 379.
 FT REPEAT 9258 9282 LRR 380.
 FT REPEAT 9282 9306 LRR 381.
 FT REPEAT 9306 9330 LRR 382.
 FT REPEAT 9330 9354 LRR 383.
 FT REPEAT 9354 9378 LRR 384.
 FT REPEAT 9378 9402 LRR 385.
 FT REPEAT 9402 9426 LRR 386.
 FT REPEAT 9426 9450 LRR 387.
 FT REPEAT 9450 9474 LRR 388.
 FT REPEAT 9474 9498 LRR 389.
 FT REPEAT 9498 9522 LRR 390.
 FT REPEAT 9522 9546 LRR 391.
 FT REPEAT 9546 9570 LRR 392.
 FT REPEAT 9570 9594 LRR 393.
 FT REPEAT 9594 9618 LRR 394.
 FT REPEAT 9618 9642 LRR 395.
 FT REPEAT 9642 9666 LRR 396.
 FT REPEAT 9666 9690 LRR 397.
 FT REPEAT 9690 9714 LRR 398.
 FT REPEAT 9714 9738 LRR 399.
 FT REPEAT 9738 9762 LRR 400.
 FT REPEAT 9762 9786 LRR 401.
 FT REPEAT 9786 9810 LRR 402.
 FT REPEAT 9810 9834 LRR 403.
 FT REPEAT 9834 9858 LRR 404.
 FT REPEAT 9858 9882 LRR 405.
 FT REPEAT 9882 9906 LRR 406.
 FT REPEAT 9906 9930 LRR 407.
 FT REPEAT 9930 9954 LRR 408.
 FT REPEAT 9954 9978 LRR 409.
 FT REPEAT 9978 10002 LRR 410.
 FT REPEAT 10002 10026 LRR 411.
 FT REPEAT 10026 10050 LRR 412.
 FT REPEAT 10050 10074 LRR 413.
 FT REPEAT 10074 10098 LRR 414.
 FT REPEAT 10098 10122 LRR 415.
 FT REPEAT 10122 10146 LRR 416.
 FT REPEAT 10146 10170 LRR 417.
 FT REPEAT 10170 10194 LRR 418.
 FT REPEAT 10194 10218 LRR 419.
 FT REPEAT 10218 10242 LRR 420.
 FT REPEAT 10242 10266 LRR 421.
 FT REPEAT 10266 10290 LRR 422.
 FT REPEAT 10290 10314 LRR 423.
 FT REPEAT 10314 10338 LRR 424.
 FT REPEAT 10338 10362 LRR 425.
 FT REPEAT 10362 10386 LRR 426.
 FT REPEAT 10386 10410 LRR 427.
 FT REPEAT 10410 10434 LRR 428.
 FT REPEAT 10434 10458 LRR 429.
 FT REPEAT 10458 10482 LRR 430.
 FT REPEAT 10482 10506 LRR 431.
 FT REPEAT 10506 10530 LRR 432.
 FT REPEAT 10530 10554 LRR 433.
 FT REPEAT 10554 10578 LRR 434.
 FT REPEAT 10578 10602 LRR 435.
 FT REPEAT 10602 10626 LRR 436.
 FT REPEAT 10626 10650 LRR 437.
 FT REPEAT 10650 10674 LRR 438.
 FT REPEAT 10674 10698 LRR 439.
 FT REPEAT 10698 10722 LRR 440.
 FT REPEAT 10722 10746 LRR 441.
 FT REPEAT 10746 10770 LRR 442.
 FT REPEAT 10770 10794 LRR 443.
 FT REPEAT 10794 10818 LRR 444.
 FT REPEAT 10818 10842 LRR 445.
 FT REPEAT 10842 10866 LRR 446.
 FT REPEAT 10866 10890 LRR 447.
 FT REPEAT 10890 10914 LRR 448.
 FT REPEAT 10914 10938 LRR 449.
 FT REPEAT 10938 10962 LRR 450.
 FT REPEAT 10962 10986 LRR 451.
 FT REPEAT 10986 11010 LRR 452.
 FT REPEAT 11010 11034 LRR 453.
 FT REPEAT 11034 11058 LRR 454.
 FT REPEAT 11058 11082 LRR 455.
 FT REPEAT 11082 11106 LRR 456.
 FT REPEAT 11106 11130 LRR 457.
 FT REPEAT 11130 11154 LRR 458.
 FT REPEAT 11154 11178 LRR 459.
 FT REPEAT 11178 11202 LRR 460.
 FT REPEAT 11202 11226 LRR 461.
 FT REPEAT 11226 11250 LRR 462.
 FT REPEAT 11250 11274 LRR 463.
 FT REPEAT 11274 11298 LRR 464.
 FT REPEAT 11298 11322 LRR 465.
 FT REPEAT 11322 11346 LRR 466.
 FT REPEAT 11346 11370 LRR 467.
 FT REPEAT 11370 11394 LRR 468.
 FT REPEAT 11394 11418 LRR 469.
 FT REPEAT 11418 11442 LRR 470.
 FT REPEAT 11442 11466 LRR 471.
 FT REPEAT 11466 11490 LRR 472.
 FT REPEAT 11490 11514 LRR 473.
 FT REPEAT 11514 11538 LRR 474.
 FT REPEAT 11538 11562 LRR 475.
 FT REPEAT 11562 11586 LRR 476.
 FT REPEAT 11586 11610 LRR 477.
 FT REPEAT 11610 11634 LRR 478.
 FT REPEAT 11634 11658 LRR 479.
 FT REPEAT 11658 11682 LRR 480.
 FT REPEAT 11682

RT progenitor cells and their myogenic descendants.";
 RL Development 119:419-431(1993).
 [2]
 RP ERRATUM.
 RX MEDLINE=94244488; PubMed=7910553;
 RA Lints T.J., Parsons L.M., Hartley L., Lyons I., Harvey R.P.;
 RL Development 119:969-969(1993).
 [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129/J;
 RX MEDLINE=98453405; PubMed=9778505;
 RA Searcy R.D., Vincent E.B., Liberatori C.M., Yutzy K.B.;
 RT "A GATA-dependent nkx-2.5 regulatory element activates early cardiac
 gene expression in transgenic mice.";
 RL Development 125:4461-4470(1998).
 [4]
 RP SEQUENCE OF 17-318 FROM N.A.
 RX MEDLINE=93376760; PubMed=7690144;
 RA Komuro I., Izumo S.;
 RT "Csk: a murine homeobox-containing gene specifically expressed in the
 developing heart.";
 RL Proc. Natl. Acad. Sci. U.S.A. 90:8145-8149(1993).
 CC -1- FUNCTION: IMPLICATED IN COMMITMENT TO AND/OR DIFFERENTIATION OF
 THE MYOCARDIAL LINEAGE.
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- TISSUE SPECIFICITY: PREDOMINANTLY IN THE ADULT AND EMBRYONIC
 HEART, AND TO A LESSER EXTENT IN LINGUAL MUSCLE, SPLEEN AND
 STOMACH.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSION PRECEDES THE ONSET OF MYOGENIC
 DIFFERENTIATION, AND CONTINUES IN CARDIOMYOCYTES OF EMBRYONIC,
 FETAL AND ADULT HEARTS. IT IS ALSO EXPRESSED LATERALLY IN FUTURE
 PHARYNGEAL ENDODERM WHICH IS BELIEVED TO PRODUCE THE HEART
 INDUCER. AFTER FOREGUT CLOSURE EXPRESSION IN ENDODERM IS LIMITED
 TO THE PHARYNGEAL FLOOR. DORSAL TO THE DEVELOPING HEART TUBE, THE
 THYROID PRIMORDIUM A DERIVATIVE OF THE PHARYNGEAL FLOOR CONTINUES
 TO EXPRESS THE PROTEIN AFTER ITS LEVELS DIMINISH IN THE REST OF
 THE PHARYNX.
 CC -1- SIMILARITY: BELONGS TO THE NK-2 HOMEBOX FAMILY.
 CC -1- SIMILARITY: CONTAINS 1 HOMEBOX DOMAIN.
 CC -1- CAUTION: REF.3 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO
 FRAMESHIFTS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL, L20300; -; NOT ANNOTATED CDS.
 DR EMBL, AF091351; AAC97934.1; -.
 DR EMBL, X75415; CAA53169.1; -.
 DR HSSP, F23441; 1FTT.
 DR TRANSFAC; T01675; -.
 DR MGD; MGI:97350; Nkx2-5.
 DR InterPro; IPR001356; Homeobox.
 DR Pfam; PF00046; homeobox; 1.
 DR PRINTS; PR00024; HOMEBOX.
 DR ProDom; PD00010; Homeobox; 1.
 DR SMART; SM00389; HOX; 1.
 DR PROSITE; PS00027; HOMEBOX_1; 1.
 DR PROSITE; PS00071; HOMEBOX_2; 1.
 KW Homeobox; DNA-binding; Developmental protein; Nuclear protein.
 FT DOMAIN 41 107 ALA/PRO-RICH.
 FT DNA_BIND 137 196 HOMEBOX.
 FT DOMAIN 206 280 ALA/PRO-RICH.
 FT CONFLICT 206 280 A -> R (IN RBF. 4).
 FT CONFLICT 164 165 AP -> PA (IN REF. 1).
 FT CONFLICT 234 234 A -> T (IN REF. 3).
 FT CONFLICT 276 277 OP -> HA (IN REF. 1).
 SQ SEQUENCE 318 AA; 34162 MW; 6A3E7859F46D4FA6 CRC64;

Query Match 7.6%; Score 79.5; DB 1; Length 318;
 Best Local Similarity 22.9%; Pred. No. 3.6;
 Matches 40; Conservative 13; Mismatches 63; Indels 59; Gaps 6;

QY 46 PNEYDV-----SEYGRBLHFRYYTDTGFCRSKAKVTEIIVC 82
 DB 11 PFSVDIINLEQOGRSLASGLSARLEATLAPASCMLAFAFPEAYSGPEAASGLAEL-- 68
 QY 83 SGQCGPARLLINAIIGRGKWMRPSGDPFRCIPDRYRAQRVQLCPG-----GEAPRARK 135
 DB 69 RAEWGPAPSPPC-----SPAPPAFTFY-----PGAYGDDPPAKDPRADK 109
 QY 136 VRLVASCKCKRLTFPHNQSELKDFETEARPPQGRKPPRARSAAKANOAELENNY 190
 DB 110 KELCALQKAVLEDKAETDG-----AERPRARRRRRRVLFQGAQYELERRR 156

Search completed: March 28, 2003, 14:16:50
 Job time : 16.6933 secs

GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 Comphen Ltd.

OM protein - protein search, using sw model

Run on: March 28, 2003, 14:12:16 ; Search time 54.72 Seconds
(without alignments)
715.441 Million cell updates/sec

Title: US-09-867-274-2

Perfect score: 1049

Sequence: 1 QGMAFKNDATETIPELGEY.....KPRPARSAKNAQLEENAY 190

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPREMBL_21:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriaph:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match length	DB	ID	Description
1	344	32.8	206	11	Q9CQN4	Q9CQN4 mus musculus
2	341	32.5	206	4	Q96HJ7	Q96HJ7 homo sapien
3	270	25.7	134	4	Q9Y3U3	Q9Y3U3 homo sapien
4	106	10.1	272	13	Q9PW80	Q9PW80 gallus galli
5	104	9.9	272	13	Q9PUK2	Q9PUK2 gallus galli
6	90.5	8.6	168	4	Q9H772	Q9H772 homo sapien
7	89.5	8.5	184	4	Q60565	Q60565 homo sapien
8	89.5	8.5	184	4	Q8WNY1	Q8WNY1 macaca mulia
9	89.5	8.5	215	8	Q9MFB6	Q9MFB6 beta vulgar
10	88.5	8.4	168	11	Q88273	Q88273 mus musculus
11	88	8.4	1114	11	Q9KXW7	Q9KXW7 mus musculus
12	87.5	8.3	184	13	Q73755	Q73755 gallus galli
13	87.5	8.3	558	4	Q96D88	Q96D88 homo sapien
14	87.5	8.3	585	4	Q9BWN1	Q9BWN1 homo sapien
15	87.5	8.3	585	4	Q8WXX2	Q8WXX2 homo sapien
16	86.5	8.2	182	13	Q73754	Q73754 xenopus lae

17	86	8.2	184	11	Q35793	Q35793 rattus norv
18	86	8.2	184	11	Q70326	Q70326 mus musculus
19	85	8.1	191	4	Q96168	Q96168 homo sapien
20	83	7.9	958	5	Q9V560	Q9V560 drosophila
21	82.5	7.9	322	4	Q96F77	Q96F77 homo sapien
22	82.5	7.9	329	2	Q9W60	Q9W60 acetobacter
23	82	7.8	141	6	Q95J85	Q95J85 monodelphis
24	82	7.8	988	6	Q97867	Q97867 sus scrofa
25	81.5	7.8	661	5	Q9V7U8	Q9V7U8 drosophila
26	81.5	7.8	669	5	Q8T0D3	Q8T0D3 drosophila
27	81	7.7	270	13	P70041	P70041 xenopus lae
28	81	7.7	1096	10	Q9X524	Q9X524 oryza sativ
29	80.5	7.7	267	4	Q95813	Q95813 homo sapien
30	80.5	7.7	715	4	Q9Y4Q3	Q9Y4Q3 homo sapien
31	80	7.6	219	12	Q91F11	Q91F11 tupia leu
32	80	7.6	258	4	Q8FAT2	Q8FAT2 homo sapien
33	80	7.6	712	4	Q9Y557	Q9Y557 homo sapien
34	80	7.6	720	4	Q9N123	Q9N123 homo sapien
35	80	7.6	720	4	Q9BRE0	Q9BRE0 homo sapien
36	80	7.6	891	10	Q8RYT6	Q8RYT6 oryza sativ
37	80	7.6	1440	5	Q20204	Q20204 caenorhabdi
38	80	7.6	1480	5	Q9V7P8	Q9V7P8 drosophila
39	80	7.6	1504	5	Q9XYV4	Q9XYV4 drosophila
40	80	7.6	1504	5	Q9V7P9	Q9V7P9 drosophila
41	79.5	7.6	318	11	Q925V3	Q925V3 mus musculus
42	79.5	7.6	386	10	Q8S5P2	Q8S5P2 oryza sativ
43	79.5	7.6	500	10	Q9M6C2	Q9M6C2 trifolium r
44	79.5	7.6	1081	4	Q76065	Q76065 homo sapien
45	79	7.5	220	10	Q9PP95	Q9PP95 oryza sativ

ALIGNMENTS

RESULT 1	Q9CQN4	PRELIMINARY;	PRT;	206 AA.
ID	Q9CQN4			
AC	Q9CQN4			
DT	01-JUN-2001 (TREMBLrel. 17, Created)			
DT	01-JUN-2001 (TREMBLrel. 17, Last sequence update)			
DT	01-MAR-2002 (TREMBLrel. 20, Last annotation update)			
DE	0610006G05RIK protein (RIKEN CDNA 0610006G05 gene).			
GN	0610006G05RIK.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.			
OX	NCBI_TaxId=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=C57BL/6J; TISSUE=PANCREAS, AND KIDNEY;			
RX	MEDLINE=21085660; PubMed=11217851;			
RA	Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,			
RA	Aizawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,			
RA	Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamada I.,			
RA	Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,			
RA	Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,			
RA	Fleischmann W., Gaasterland T., Gissi C., King B., Kochia H.,			
RA	Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,			
RA	Schriml L.M., Stauber F., Suzuki R., Tomita M., Wagner L., Washio T.,			
RA	Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barns G.,			
RA	Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,			
RA	Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,			
RA	Guerninich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,			
RA	Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,			
RA	Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,			
RA	Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,			
RA	Suzuki H., Toyooka K., Wang K.H., Wetz C., Whitaker C., Wilming L.,			
RA	Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohatsu S.,			
RA	Hayashizaki Y.,			
RT	"Functional annotation of a full-length mouse cDNA collection.";			
RL	Nature 403:685-690(2001).			
RN	[2]			
RP	SEQUENCE FROM N.A.			

RC TISSUE=KIDNEY;
RA Strausberg R;
PL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases
DR EMBL; AK007967; BAB25378.1; -;
DR EMBL; AK002240; BAB2157.1; -;
DR EMBL; AK002396; BAB2268.1; -;
DR EMBL; AK007893; BAB2533.1; -;
DR EMBL; BC021458; AAH21458.1; -;
DR MGI; MGI:1913292; 0610006G05R1k.
DR InterPro: IPR000359; Cys_knot.
DR PROSITE; PS01225; CTCK_2; 1.
SQ SEQUENCE 206 AA; 23174 MD; 70D24819EB06CB9C CEC64;

Query Match	32.8%;	Score 344;	DB 11;	Length 206;
Best Local Similarity	42.7%;	Pred. No. 1.4e-27;		
Matches 79;	Conservative 30;	Mismatches 62;	Indels 14;	Gaps 8;

[illegible]

```

Oy 117 PAORVOLLCPGEAPPAKRVLASCKKRLTRPHNOSLKDFTGEAPRPOCKGRKPPRA 176
    | | | | | : : : | | | | | : : : |
Db 139 RTORQLQCGDG-STRTYKITVTACKCKRYTRQNESSNPFESVPAKPAQHREKRA 197

```

Qy	177	-RS	AK	180
	:	:	:	
Db	198	SK	SK	202

RESULT 2
Q96HJ7
ID Q96HJ7 PRELIMINARY; PRT; 206 AA.

DT 01-DEC-2001 (Tremblé1. 19, Created)
DT 01-DEC-2001 (Tremblé1. 19, last sequence update)
DT 01-MAR-2002 (Tremblé1. 20, last annotation update)
DE Similar to RIKEN cDNA 0610006G05 gene (CDA019).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;

RP SEQUENCE FROM N.A.
RC TISSUE=BONE MARROW;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases

RA Liu F., Xu X.R., Qian B.Z., Xiao H., Chen Z., Han Z.;
RT "A novel gene expressed in human pheochromocytoma";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases

DR EMBL; AF361494; AAL57219.1; -
DR InterPro; IPR000359; Cys_knot.
DR PROSITE; PS01225; CTKC 2; 1.
SQ SEQUENCE 206 AA; 23306 MW; 9FB3CC41EAB53B34 CRC64

Query Match	32.5%	Score	341	DB	4	Length	206
Best Local Similarity	42.7%	Pred. No.	2.8e-27				
Matches	79	Conservative	30	Mismatches	62	Indels	14
						Gaps	8

```
Oy      5 AFKDATEIIPELGEYEP-PELENNKTMRANGRPRHPPETTKDVS---EYSREL 60
          ||||| : | : | : | : | : | : | : | : | : | : | : | : |
Db      23 AFKNDATEIL--YSHVVKPVPAHPSNSTLQANRGR--HFSNTEGLDRNRVRQVGREL 78
```

```
Oy      61 HFTRYVTDGPSCSAKPYTELVCSGQCOPARLLPNAIGRG---KWM-RPSGPDFRCIPDRY 116
        |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:
Db      79 RSTKYISDGQCTSI SPLKELVCA GECPLPVLPWNIGGGGT KYMRSRRSSQEMRCVN DKT 138
```

[illegible]

RESULT 3	
Q9Y3U3	
ID Q9Y3U3	PRELIMINARY; PRT; 134 AA

DT 01-NOV-1999 (TRENMBLrel. 12, Created)
DT 01-NOV-1999 (TRENMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TRENMBLrel. 19, Last annotation update)
DE Hypothetical 15.3 KDa protein (Fragment).
GN DKFZ564D206.

OS *Onchocerca* spp.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi
 CC Mammalia; Eutheria; Primates; Cactariini; Hominae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.

RA Warbitt R., Heubner D., Mewes H.W., Gassenhuber J., Wiemann S.;
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.

DR InterPro; IPR000359; Cys_knot.
DR PROSITE; PS01225; CTCK_2; 1.
KW Hypothetical protein.

Query Match	25.7%;	Score 270;	DB 4;	Length 134;
Best Local Similarity	44.2%;	Pred. No. 3.7e-20;		
Matches 57;	Conservative 23;	Mismatches 43;	Indels 6;	Gaps 4;

```
Qy      57 CRELHFRITYTDGPECRSAKPYTELVCGGCCGPALLLPNAIGR---KW-RPSGPDFRCI 112
        |||:::|||||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
Db      3 CRELRSTKYISDGCCTSIPLKEIVCAGECLLLPVLPMNIGGYGTKYWRSSRGEMRCV 62
```

```

Qy      113 PDRRAQRVOLLCPGSEAPRAKRVLASCKKRLTRFPNOSLKDFGTEARPOQGRKP 172
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      63 NDKRTQRIQLQCQGD-STRYKITYVACKCKCRYTRQNESSHNFEPSMAKPVQHRE 121

```

QY	173	RPR-RSAK	180
		: :	
Db	122	RKRASKSSK	130

RESULT 4	
Q9PWBO	
ID Q9PWBO	PRELIMINARY;
Q9PWBO	PRT; 272 AA

DT	01-MAY-2000 (TEMBLrel. 13, Created)
DT	01-MAY-2000 (TEMBLrel. 13, last sequence update)
DT	01-JUN-2002 (TEMBLrel. 21, last annotation update)
DE	Cerberus homolog.

0S Gallus gallus (chicken).
0C Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi;
0C Actinoptera, Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
0C Gallus.

RN [1] -
RP SEQUENCE FROM N.A.
RX MEDLINE=99439862; PubMed=10508582;

RA	Levin M.;
RT	"Cerberus regulates left-right asymmetry of the embryonic head and
RT	heart.";

RL Curr. Biol. 9:931-938(1999).
 DR EMBL; AF139721; AAD51610.1; -
 DR InterPro; IPR000359; Cys_knot.
 DR InterPro; IPR004133; DAN_dom.
 DR Pfam; PF03045; DAN; 1.
 DR SMART; SM00041; CT; 1.
 DR PROSITE; PS01225; CTCK_2; 1.
 DR PROSITE; PS01225; CTCK_2; 1.
 SQ SEQUENCE 272 AA; 31224 MW; 36E4C9F719711BCA CRC64;

Query Match 10.1%; Score 106; DB 13; Length 272;
 Best Local Similarity 23.7%; Pred. No. 0.0066;
 Matches 32; Conservative 20; Mismatches 57; Indels 26; Gaps 5;

QY 22 EPPPELENNK---TWNRALNGRPPHHPETKDVSEYSCRELAFTRYVTDGCRSAKPV 77
 DB 129 EPPYRKDAKKFMDHFLRKNSASAEVVLPIKTNEMHGETCRTLPSQSVAHSCCKV-IV 187
 QY 78 TELVSGGCGPARLLPNAIGRGKMRPSPGDFR-----CIPDRYRAQRVQLCPGGGA 130
 DB 188 QNNLCGFGKSS-----FHVPGPDDRLYTFCKSLPTKFSMGLDINCT-SSV 233
 QY 131 PRARKRLVASCKCK 145
 DB 234 PVKAKMIVECNCE 248

RESULT 5

Q9PUK2 PRELIMINARY; PRT; 272 AA.

AC Q9PUK2;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Caronte.
 GN CAR.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OK NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Rodriguez Esteban C., Capdevila J., Economides A.N., Pascual J.,
 RA Ortiz A., Izpisua Belmonte J.C.;
 RT "Caronte, a novel cer-like protein, mediates the establishment of
 RT embryonic left-right asymmetry.";
 RL Nature 0:0-0(1999).
 DR EMBL; AF179484; AAD55581.1; -
 DR InterPro; IPR000359; Cys_knot.
 DR InterPro; IPR004133; DAN_dom.
 DR Pfam; PF03045; DAN; 1.
 DR SMART; SM00041; CT; 1.
 DR PROSITE; PS01225; CTCK_2; 1.
 DR PROSITE; PS01225; CTCK_2; 1.
 SQ SEQUENCE 272 AA; 31201 MW; 0DB9729715771BC2 CRC64;

Query Match 9.9%; Score 104; DB 13; Length 272;
 Best Local Similarity 23.7%; Pred. No. 0.011;
 Matches 32; Conservative 19; Mismatches 58; Indels 26; Gaps 5;

QY 22 EPPPELENNK---TWNRALNGRPPHHPETKDVSEYSCRELAFTRYVTDGCRSAKPV 77
 DB 129 EPPYRKDAKKFMDHFLRKNSASAEVVLPIKTNEMHGETCRTLPSQSVAHSCCKV-IV 187
 QY 78 TELVSGGCGPARLLPNAIGRGKMRPSPGDFR-----CIPDRYRAQRVQLCPGGGA 130
 DB 188 QNNLCGFGKSS-----FHVPGPDDRLYTFCKSLPTKFSMGLDINCT-SSV 233
 QY 131 PRARKRLVASCKCK 145
 DB 234 PVKAKMIVECNCE 248

RESULT 6

Q9H772 PRELIMINARY; PRT; 168 AA.

AC Q9H772;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE CDNA: FLJ21195 f18, clone COL00185.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OK NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=COLON;
 RA Kawabata A., Hiki T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
 RA Okitani R., Ota T., Suzuki Y., Obayashi M., Nishi T., Shibahara T.,
 RA Tanaka T., Nakamura Y., Iisgai T., Sugano S.;
 RT "NEDO human cDNA sequencing project.";
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK024848; BAB15026.1; -
 DR InterPro; IPR000359; Cys_knot.
 DR InterPro; IPR004133; DAN_dom.
 DR Pfam; PF03045; DAN; 1.
 DR SMART; SM00041; CT; 1.
 DR PROSITE; PS01225; CTCK_2; 1.
 DR PROSITE; PS01225; CTCK_2; 1.
 SQ SEQUENCE 168 AA; 19320 MW; D5A4E818BF8C0E CRC64;

Query Match 8.6%; Score 90.5; DB 4; Length 168;
 Best Local Similarity 26.5%; Pred. No. 0.16;
 Matches 27; Conservative 19; Mismatches 43; Indels 13; Gaps 5;

QY 48 ETKVSEYSCRELAFTRYVTDGCRSAKPVTELVCSSGCGPARLLPNAIGRGKMRPSPG 107
 DB 64 ERKYLSDWCKTQPLRGVSESGRS--RTILNRCYQGN-SFIPPHV-----KKEE 115
 QY 108 DFR---CIPDRYRAQRVQLCPGGGA-PRARKRLVASCKCK 144
 DB 116 SFOSCAFCKPQRTVSVVLECPGLDPPFLKTIQVKQCRG 157

RESULT 7

Q60565 PRELIMINARY; PRT; 184 AA.

ID Q60565;
 AC Q60565;
 DT 01-AUG-1998 (TREMBLrel. 07, Created)
 DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE GREMLIN (DRM).
 GN DRM.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OK NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98325381; PubMed=9660951;
 RA Hau D.R., Economides A.N., Wang X., Eimon P.M., Harland R.M.;
 RT "The Xenopus dorsalizing factor Gremlin identifies a novel family of
 RT secreted proteins that antagonize BMP activities.";
 RL Mol. Cell 1:673-683(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Tate G., Mitsuya T.;
 RT "Human Gremlin homologue";
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA McMahon R.A., Murphy M., Clarkson M.R., Godson C., Martin F.,
 RA Brady H.R.;
 RT "HIG-2, a mesangial cell gene differentially induced in high glucose,
 RT is human gremlin.";
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 RN [4]

RP	SEQUENCE FROM N.A.
RC	TISSUE=SMALL_INTESTINE;
RX	MEDLINE=2035496; PubMed=10894942;
RA	Topol I.Z., Modi W.S., Koochekpour S., Blair D.G.;
RT	"DRN/GEMBLIN (CKTSFBI) maps to human chromosome 15 and is highly
RT	expressed in adult and fetal brain.";
RL	Cytogenet. Cell Genet. 89:79-84(2000).
DR	EMBL; AF045800; AAC39725.1; -
DR	EMBL; AB032337; BAA84462.1; -
DR	EMBL; AF110137; AAF06677.1; -
DR	EMBL; AF154054; AAG23891.1; -
DR	InterPro; IPRO00359; Cys_knot.
DR	InterPro; IPRO04133; DAN_dom.
DR	Pfam; PF03045; DAN; 1.
DR	SMART; SM00041; CT; 1.
SQ	SEQUENCE 184 AA; 20697 MW; 4B588598DE12C47E CRC64;
Query Match	8.5%; Score 89.5; DB 4; Length 184;
Best Local Similarity	24.1%; Pred. No. 0.22; Mismatches 57; Indels 41; Gaps
Matches 38; Conservative 22;	
OY	18 GEYPEPPELENNK-----TMAENGGRRPHHPFTKDVSEYSCELIHFT--R 64
Dd	31 GAIP-PDPAQHNDSEQTSPQQPGSRNRGRGGRTAMPGE--EVLESSQEALHTERK 87
OY	65 YVTDGPCR-----SAKPYTELVCSGCGPARLLPNAIGRGKWRPSGDPR- 110
Dd	88 YLRDWCKTKPLKOTIHEGCNSRTIINFVGYQC-N-SFYIPRH-----RKESGFOS 140
OY	111 ---CIPIIDRYAQRVOLLCGSGEAP-RARKRVLYASCKC 144
Dd	141 CSFCPKPKKTMMVTILNCPELOPTTKKKVTRVKQCR 178
RESULT 8	
OSMNY1	PRELIMINARY; PRT; 184 AA.
ID	OSMNY1
AC	Q8WNY1
DT	01-MAR-2002 (TREMBLrel. 20, Created)
DT	01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE	Grenlin.
OS	Macaca mulatta (rhesus macaque).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC	Cercopitheciinae; Macaca.
OX	NCBI_TaxId=9544;
RN	[1]
RP	SEQUENCE FROM N.A.
RA	Christenson L.K., Duffy D.M.;
RL	Submitted (OCT-2001) to the EMBL/genBank/DBJ databases.
DR	EMBL; AF439783; AAL32022.1; -
DR	InterPro; IPRO00359; Cys_knot.
DR	InterPro; IPRO04133; DAN_dom.
DR	Pfam; PF03045; DAN; 1.
DR	SMART; SM00041; CT; 1.
SQ	SEQUENCE 184 AA; 20697 MW; 4B588598DE12C47E CRC64;
Query Match	8.5%; Score 89.5; DB 6; Length 184;
Best Local Similarity	24.1%; Pred. No. 0.22; Mismatches 57; Indels 41; Gaps
Matches 38; Conservative 22;	
OY	18 GEYPEPPELENNK-----TMAENGGRRPHHPFTKDVSEYSCELIHFT--R 64
Dd	31 GAIP-PDPAQHNDSEQTSPQQPGSRNRGRGGRTAMPGE--EVLESSQEALHTERK 87
OY	65 YVTDGPCR-----SAKPYTELVCSGCGPARLLPNAIGRGKWRPSGDPR- 110
Dd	88 YLRDWCKTKPLKOTIHEGCNSRTIINFVGYQC-N-SFYIPRH-----RKESGFOS 140
OY	111 ---CIPIIDRYAQRVOLLCGSGEAP-RARKRVLYASCKC 144
Dd	141 CSFCPKPKKTMMVTILNCPELOPTTKKKVTRVKQCR 178

ID	NAME	PRELIMINARY	PRT	215 AA
AC	Q9MFB6			
DT	01-OCT-2000 (TREMBlrel. 15, Created)			
DT	01-OCT-2000 (TREMBlrel. 15, Last sequence update)			
DT	01-JUN-2001 (TREMBlrel. 17, Last annotation update)			
DE	Orf215 protein.			
GN	ORF215			
OS	Beta vulgaris (sugar beet).			
OC	Mitochondrion.			
OC	Eukaryota, Viridiplantae, Streptophyta, Embryophyta, Tracheophyta,			
OC	Spermatophyta, Magnoliophyta, eudicotyledons, core eudicots;			
OC	Caryophyllidae, Caryophyllales, Chenopodiaceae, Beta.			
OX	NCBI_TaxID=3555;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=TK81-O;			
FX	MEDLINE=20330382; PubMed=10871408;			
RA	Kubo T., Nishizawa S., Sugawara A., Ichihara N., Estiati A., Mikami T.;			
RT	"The complete nucleotide sequence of the mitochondrial gene of sugar			
RL	beet (<i>Beta vulgaris</i> L.) reveals a novel gene for tRNA ^{Cys} (GCA).";			
DR	Nucleic Acids Res. 28:2571-2576(2000).			
DR	EMBL; AF000396; BAA9326.1; -.			
DR	Mitochondrion.			
SO	SEQUENCE 215 AA; 24100 MW; C144AE9AE41E5E06 CRC64;			
Query Match	8.5%; Score 89.5; DB 8; Length 215;			
Best Local Similarity	28.7%; Pred. No. 0.26; Mismatches 32; Indels 31; Gaps			
Matches	31; Conservative 14; Mismatches 32; Indels 31; Gaps			
OY	3 WQAFKNDATETIIPELGXYEPPPELENNKTNRAENGRRPHPEPTKDVSEYSCRELIHF 62			
DB	134 WRGYLNDSE-----PEGGDVAVSAENYPAEN--PGHNEAAPAP--NCRSIH- 178			
OY	63 TRYVTDGPRSAKPYTELVCSCGCGPARLLPN-AIGRGK-WRPSGP 107			
DB	179 -----PCNG--PWTREPC-GNC--FLINMWTIKTGRMSWKGLP 213			
RESULT 10				
088273				
ID	088273	PRELIMINARY;	PRT;	168 AA.
AC	088273;			
DT	01-NOV-1998 (TREMBlrel. 08, Created)			
DT	01-NOV-1998 (TREMBlrel. 08, Last sequence update)			
DT	01-JUN-2002 (TREMBlrel. 21, Last annotation update)			
DE	PRDC.			
GN	PRDC OR PRDC.			
OS	Mus musculus (Mouse).			
OC	Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi;			
OC	Mammalia, Eutheria, Rodentia, Sciurognathi, Muridae, Murinae, Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=CS7BL6; TISSUE=CNS.			
FX	MEDLINE=96301233; PubMed=9639362.			
RA	Mihabe-Saegusa C., Saegusa H., Tsukahara M., Noguchi S.;			
RT	"Sequence and expression of a novel mouse gene PDC (protein related			
RL	to DAN and cerberus) identified by a gene trap approach.";			
DR	Dev. Growth Differ. 40:343-353(1998).			
DR	EMBL; AB011030; BAA29038.1; -.			
DR	MED; MGI:1344367; Prdc.			
DR	InterPro; IPR000359; Cys_knot.			
DR	InterPro; IPR004133; DAN_dom.			
DR	Pfam; PF03045; DAN; 1.			
DR	SMART; SM00041; CT; 1.			
DR	PROSITE; PS01225; CTCK 2; 1.			
SO	SEQUENCE 168 AA; 19334 MW; 6361C1561D49C281 CRC64;			

Best Local Similarity 26.5%; Pred. No. 0.25;
Matches 27; Conservative 19; Mismatches 43; Indels 13; Gaps 5;

QY 48 ETDVSEYSGREHFTRYVTDGCRSAKPTVELVCGGCGPARLLPNAIGRGKMWPSGP 107
DB 64 ERYKLSKDWCKTPTLQVSEEGCRS-RTILNRFYCGQCN-SFYIPRHV-----KKEED 115
QY 108 DFR----CIPDRYARQVOLLCPGSEAP-PARKYRLVASCKC 144
DB 116 SFQSCAFCKRQRTSVIVLECEGLDPPFRKIKIKYKRC 157

RESULT 11

09JWK7

PRELIMINARY; PRT; 1114 AA.

AC 09JWK7; PRELIMINARY; PRT; 1114 AA.
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DE Cysteine-rich protein NFX-1.
GN NFX1 OR 1300017N15RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.

RA Arlotta P., Jenkins N.A., Jay G., Ono S.J.;
RT "Isolation of a Full-length Murine NFX-1 cDNA, its Chromosomal
Location and Developmental Expression."
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF223576; AAF34700.1; -
DR MGI; 1921414; NFX1.
DR InterPro; IPR001374; R3H.
DR InterPro; IPR000967; Znf_NFX1.
DR InterPro; IPR001965; Znf_PHD.
DR InterPro; IPR001841; Znf_Ring.
DR Pfam; PF01424; R3H; 1.
DR Pfam; PF01422; ZF-NFX1; 8.
DR SMART; SM00393; R3H; 1.
DR SMART; SM00184; RING; 1.
DR SMART; SM00438; ZNF_NFX; 9.
SQ SEQUENCE 1114 AA; 123878 MW; 857C90260D28EA4B CRC64;

Query Match 8.4%; Score 88; DB 11; Length 1114;
Best Local Similarity 24.5%; Pred. No. 2.2; Indels 34; Gaps 9;

Matches 46; Conservative 22; Mismatches 86; Indels 34; Gaps 9;

QY 1 QGWAQFND-----ATEIIPELGEYPPPELENNKTNRAENGGR-PHHPFETKDVSEY 55
DB 101 QPQKANEHGNRKNKKAQGLSEGTSTSLF---SVARESGTNPREHSPSE----- 152
QY 56 SCREHFTRYVTDGCRSAKPTVELVCGGCGPARLLPNAIG--RGKMWPSGPDFRC-1 112
DB 153 -----KEVVDLPRGAKPRKKAQLTINYGRG--PKAKGRLRSEGNRMSPKSDEI 201
QY 113 PDYRAQRVOLLCPGSEAPARKYRLVASCKCKRLTFHNSQLKDGCTAASPOKGRKP 172
DB 202 PDYRAQRVOLLCPGSEAPARKYRLVASCKCKRLTFHNSQLKDGCTAASPOKGRKP 255
QY 173 RPRASAK 180
DB 256 -PKQESQR 262

RESULT 12

073755

PRELIMINARY; PRT; 184 AA.

AC 073755; PRELIMINARY; PRT; 184 AA.
DT 01-AUG-1998 (TREMBlrel. 07, Created)
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Gremilin.

OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauvia; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.

RX MEDLINE=98325381; PubMed=9660951;
RA Hsu D.R., Economides A.N., Wang X., Elmon P.M., Harland R.M.;
RT "The Xenopus dorsalizing factor Gremilin identifies a new family of
secreted proteins that antagonize BMP activities";
RL Mol. Cell 1:673-683(1998).
DR EMBL; AF045799; AAC41280.1; -
DR InterPro; IPR000359; Cys_knot.
DR InterPro; IPR004133; DAN_dom.
DR Pfam; PF03045; DAN; 1.
DR SMART; SM00041; CT; 1.
SQ SEQUENCE 184 AA; 21166 MW; 3510B448BFDD5EA CRC64;

Query Match 8.3%; Score 87.5; DB 13; Length 184;
Best Local Similarity 22.4%; Pred. No. 0.35;

Matches 34; Conservative 24; Mismatches 55; Indels 39; Gaps 8;

QY 23 PPELE-----NNKTNRAENGGRPPHHPFETKDVSEYSGREHFT--RYVTDGP 70
DB 36 PPPDKQPDSEOMOTQOOSGRHREGRGKTSMPAE--EVLESQELALITERKYLKRDW 93
QY 71 CR-----SAPVTELVCGGCGPARLLPNAIGRGKMWPSGPDFR-----CIP 113
DB 94 CTPQPKQTHHEGNSRTIIRNFYCGQN-SFYIPRHV-----RKEGSGFCSFKCP 146
QY 114 DRYRAQRVOLLCPGSEAPARKYRLVASCKC 144
DB 147 KFTMTVTWLNCPQLPPPRKKRITRVKRC 178

RESULT 13

096D88

PRELIMINARY; PRT; 558 AA.

AC 096D88; PRELIMINARY; PRT; 558 AA.
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Similar to hypothetical protein MG3121 (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.

RC TISSUE=CERVIX;
RA Strauberg R.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC010232; AAH10232.1; -
KW Hypothetical protein.
FT NON TER 1

SQ SEQUENCE 558 AA; 60917 MW; FD48933F9F3DB5B6 CRC64;

Query Match 8.3%; Score 87.5; DB 4; Length 558;
Best Local Similarity 23.2%; Pred. No. 1.2;

Matches 46; Conservative 14; Mismatches 45; Indels 93; Gaps 10;

QY 21 PPEPELENNKTNRAENGGR-----RPPHHPFETKDVSEYSGR 58
DB 249 PSPMKLEIKIAISAEQGAEGTASVSPRPRIROWTODINTPALLPKPSLGSGYSCP 308
QY 59 ELHFTRYVTDGCRSAKPTVELVCGGCGPARLLPNAIGRGKMWPSGPDFRCIPDRYRA 118
DB 309 DL-----GPPGPGCT-----WPPAPFQ-----PSRPP 332
QY 119 QRVOLLCPGSEAPRA-----RK-----VRLVASCKCKRLTFHNSQLKDF 159
DB 333 RRTTV--GGGEMARAPPPPRPCLRKVFPPLGGVGSPLTTSSTASTSF----- 381

QY 160 GTEARPO---KGRKR 173
DB 382 -SEPAEPRLGSTKGKPR 398

RESULT 14

Q9BWN1 PRELIMINARY; PRT; 585 AA.
AC Q9BWN1;
DT 01-JUN-2001 (TREMblrel. 17, Created)
DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE Hypothetical 64.3 kDa protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCB1_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=EYE;
RA Strausberg R.;
RL Submitted (NCV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC000119; AAH00119.1; -
DR InterPro: IPR002965; P_rich_extensn.
DR PRINTS; PRO1217; PRICHEXTENSN.
KW Hypothetical protein.
SQ SEQUENCE 585 AA; 64328 MW; EE706853CE2D2742 CRC64;

Query March 8.3%; Score 87.5; DB 4; Length 585;
Best Local Similarity 23.2%; Pred. No. 1.2;
Matches 46; Conservative 14; Mismatches 45; Indels 93; Gaps 10;

QY 21 PEPPPELENNKTNNRAENG-----RPPHPFETKDV-----EYSCR 58
DB 276 PSPPMKLELKIAISEAEQGAAGTASVSPRPPIQWRTQDHTPALPKPSLGSTSCP 335
QY 59 ELHFTRYVTDPGCRSAKPYTELVCSGQCGPARLLPNAIGRGKMWPPSGDPDFRCIPDRYRA 118
DB 336 DL-----GPPGPGTCT-----WPPAPQ-----PSRRL 359
QY 119 QRVQLCPGGEAPRA-----RK-----VRLVASCCKRLTRFNOSELKDF 159
DB 360 RRTTV--GGGEWAPAPPPLRKEVPLGAGVSPSLTTSSTASTSF----- 408
QY 160 GTEARPO---KGRKR 173
DB 409 -SEPAEPRLGSTKGKPR 425

RESULT 15

Q8WTK2 PRELIMINARY; PRT; 585 AA.
AC Q8WTK2;
DT 01-MAR-2002 (TREMblrel. 20, Created)
DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)
DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
DE Hypothetical 64.3 kDa protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCB1_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=RETINA;
RA Strausberg R.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC021934; AAH21934.1; -
KW Hypothetical protein.
SQ SEQUENCE 585 AA; 64344 MW; 2DFB54B1551CE4C CRC64;

Query Match 8.3%; Score 87.5; DB 4; Length 585;
Best Local Similarity 23.2%; Pred. No. 1.2;

Matches 46; Conservative 14; Mismatches 45; Indels 93; Gaps 10;
QY 21 PEPPPELENNKTNNRAENG-----RPPHPFETKDV-----EYSCR 58
DB 276 PSPPMKLELKIAISEAEQGAAGTASVSPRPPIQWRTQDHTPALPKPSLGSTSCP 335
QY 59 ELHFTRYVTDPGCRSAKPYTELVCSGQCGPARLLPNAIGRGKMWPPSGDPDFRCIPDRYRA 118
DB 336 DL-----GPPGPGTCT-----WPPAPQ-----PSRRL 359
QY 119 QRVQLCPGGEAPRA-----RK-----VRLVASCCKRLTRFNOSELKDF 159
DB 360 RRTTV--GGGEWAPAPPPLRKEVPLGAGVSPSLTTSSTASTSF----- 408
QY 160 GTEARPO---KGRKR 173
DB 409 -SEPAEPRLGSTKGKPR 425

Search completed: March 28, 2003, 14:18:46
Job time : 57.72 secs

GenCore version 5.1.4 p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 28, 2003, 14:16:22 ; Search time 12.16 Seconds
(without alignments)
917.557 Million cell updates/sec

Title: US-09-867-274-2

Perfect score: 1049

Sequence: 1 QGMQAFKNDATRIIPELGEY.....KPRPARSAKANOAELENNY 190

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 237916 seqs, 58723674 residues

Total number of hits satisfying chosen parameters: 237916

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US05_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1049	100.0	190	10	US-09-867-274-2
2	1049	100.0	213	10	US-09-867-274-5
3	932.5	88.9	185	10	US-09-867-274-4
4	932.5	88.9	208	10	US-09-867-274-6
5	758	72.3	139	10	US-09-864-761-47109
6	344	32.8	206	9	US-09-866-050A-159
7	344	32.8	206	9	US-09-866-050A-286
8	341	32.5	206	10	US-09-853-625B-2
9	341	32.5	206	10	US-09-853-625B-20
10	337	32.1	183	10	US-09-867-274-25
11	98	9.3	337	10	US-09-814-777A-126
12	89.5	8.5	184	9	US-10-044-716-10
13	89.5	8.5	184	9	US-10-159-749-2
14	87.5	8.3	184	9	US-10-159-749-6
15	86.5	8.2	182	9	US-10-159-749-4
16	86	8.2	184	9	US-10-159-749-9
17	85	8.1	181	9	US-10-159-749-8
18	82.5	7.9	462	10	US-09-814-777A-99
19	82	7.8	150	10	US-09-925-301-1324

20	81	7.7	270	9	US-09-903-170C-1	Sequence 1, Appl1
21	81	7.7	270	10	US-09-903-180B-1	Sequence 1, Appl1
22	81	7.7	270	10	US-09-903-187A-1	Sequence 1, Appl1
23	81	7.7	270	10	US-09-903-471A-1	Sequence 1, Appl1
24	81	7.7	270	10	US-09-903-188A-1	Sequence 1, Appl1
25	81	7.7	270	10	US-09-903-323A-1	Sequence 1, Appl1
26	81	7.7	270	10	US-09-903-325A-1	Sequence 1, Appl1
27	80.5	7.7	267	9	US-10-044-716-12	Sequence 12, Appl1
28	80.5	7.7	267	9	US-09-089-818B-8	Sequence 8, Appl1
29	80	7.6	1480	12	US-10-011-064-5	Sequence 5, Appl1
30	79	7.5	336	10	US-09-756-186-8	Sequence 8, Appl1
31	78	7.4	1210	10	US-09-725-433-2	Sequence 2, Appl1
32	77	7.3	510	10	US-09-866-562-56	Sequence 56, Appl1
33	76.5	7.3	302	9	US-10-029-180-10	Sequence 10, Appl1
34	76	7.2	671	10	US-09-925-301-844	Sequence 844, Appl1
35	75.5	7.2	1436	9	US-10-042-431-78	Sequence 78, Appl1
36	75.5	7.2	1436	9	US-09-759-130B-448	Sequence 48, Appl1
37	75	7.1	4123	9	US-10-213-509-5	Sequence 5, Appl1
38	73.5	7.0	400	9	US-10-026-021-5	Sequence 249, Appl1
39	73.5	7.0	685	10	US-09-771-161A-249	Sequence 250, Appl1
40	73.5	7.0	685	10	US-09-771-161A-250	Sequence 251, Appl1
41	73.5	7.0	685	10	US-09-771-161A-251	Sequence 250, Appl1
42	72.5	6.9	104	10	US-09-764-887-260	Sequence 260, Appl1
43	72.5	6.9	247	9	US-10-050-704-275	Sequence 275, Appl1
44	72.5	6.9	547	9	US-09-977-418-4	Sequence 4, Appl1
45	72	6.9	251	10	US-09-764-853-522	Sequence 522, Appl1

ALIGNMENTS

RESULT 1
US-09-867-274-2
Sequence 2, Application US/09867274
; Patent No. US2002010650A1
; GENERAL INFORMATION:
; APPLICANT: Paszty, Christopher
; APPLICANT: Gao, Yongming
; TITLE OF INVENTION: Cysteine knot Polypeptides: Cloaked-2 Molecules and Uses Thereof
; FILE REFERENCE: 01017/37428
; CURRENT APPLICATION NUMBER: US/09/867,274
; PRIOR FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: US 60/208,550
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: US 60/223,542
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 190
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-867-274-2

Query Match 100.0%; Score 1049; DB 10; Length 190;
Best Local Similarity 100.0%; Pred. No. 4.8e-96;
Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QGMQAFKNDATRIIPELGEYPPPELENNKTWNAENGRRPDPHPFETKOVSEYSCREL 60
DB 1 QGMQAFKNDATRIIPELGEYPPPELENNKTWNAENGRRPDPHPFETKOVSEYSCREL 60
QY HHTRYVTGDPCKSAPVVELVCSGGCPARLLPNAIGKWWRPSPGPPRCIPDYRROR 120
DB HHTRYVTGDPCKSAPVVELVCSGGCPARLLPNAIGKWWRPSPGPPRCIPDYRROR 120
QY 121 VOLLCPGGAAPPARVRLVASCKCKRLTRFNQSELKDFGTEAARPOKGRKPRPARSAK 180
DB 121 VOLLCPGGAAPPARVRLVASCKCKRLTRFNQSELKDFGTEAARPOKGRKPRPARSAK 180
QY 181 ANOAELENNY 190
DB 181 ANOAELENNY 190

RESULT 2

US-09-867-274-5
 ; Sequence 5, Application US/09867274
 ; Patent No. US20020106650A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Paszty, Christopher
 ; APPLICANT: Gao, Yongming
 ; TITLE OF INVENTION: Cysteine Knot Polypeptides: Cloaked-2 Molecules and Uses Thereof
 ; FILE REFERENCE: 01017/37428
 ; CURRENT APPLICATION NUMBER: US/09/867,274
 ; PRIOR FILING DATE: 2001-05-29
 ; PRIOR APPLICATION NUMBER: US 60/208,550
 ; PRIOR FILING DATE: 2000-06-01
 ; PRIOR APPLICATION NUMBER: US 60/223,542
 ; PRIOR FILING DATE: 2000-08-04
 ; NUMBER OF SEQ ID NOS: 25
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 5
 ; LENGTH: 213
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-867-274-5

Query Match 100.0%; Score 1049; DB 10; Length 213;
 Best Local Similarity 100.0%; Pred. No. 5.5e-96;
 Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 OGQWAFKNDATETIIPELGEYEPPEPELENNKTNRAENGRRPHHPETKDVSEYSCREL 60
 DB 24 OGQWAFKNDATETIIPELGEYEPPEPELENNKTNRAENGRRPHHPETKDVSEYSCREL 83
 QY 61 HFTRYVTDGCRSAKPVTETVCSGCCGPARRLLPVAIGRGKWRPSPGDFPCIPDRYRAOR 120
 DB 84 HFTRYVTDGCRSAKPVTETVCSGCCGPARRLLPVAIGRGKWRPSPGDFPCIPDRYRAOR 143
 QY 121 VOLLCPGGEAPRAKRVLVASCKCKRLTRFNOSSELDGFTGEARPOKGRKPRRARSAR 180
 DB 144 VOLLCPGGEAPRAKRVLVASCKCKRLTRFNOSSELDGFTGEARPOKGRKPRRARSAR 203
 QY 181 ANQAELENAY 190
 DB 204 ANQAELENAY 213

RESULT 3

US-09-867-274-4
 ; Sequence 4, Application US/09867274
 ; Patent No. US20020106650A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Paszty, Christopher
 ; APPLICANT: Gao, Yongming
 ; TITLE OF INVENTION: Cysteine Knot Polypeptides: Cloaked-2 Molecules and Uses Thereof
 ; FILE REFERENCE: 01017/37428
 ; CURRENT APPLICATION NUMBER: US/09/867,274
 ; PRIOR FILING DATE: 2001-05-29
 ; PRIOR APPLICATION NUMBER: US 60/208,550
 ; PRIOR FILING DATE: 2000-06-01
 ; PRIOR APPLICATION NUMBER: US 60/223,542
 ; PRIOR FILING DATE: 2000-08-04
 ; NUMBER OF SEQ ID NOS: 25
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 4
 ; LENGTH: 185
 ; TYPE: PRT
 ; ORGANISM: Mus musculus
 US-09-867-274-4

Query Match 88.9%; Score 932.5; DB 10; Length 185;
 Best Local Similarity 88.4%; Pred. No. 1.4e-84;
 Matches 168; Conservative 10; Mismatches 7; Indels 5; Gaps 2;

QY 1 OGQWAFKNDATETIIPELGEYEPPEPELENNKTNRAENGRRPHHPETKDVSEYSCREL 60
 DB 1 OGQWAFKNDATETIIPELGEYEPPEPELENNKTNRAENGRRPHHPETKDVSEYSCREL 58
 QY 61 HFTRYVTDGCRSAKPVTETVCSGCCGPARRLLPVAIGRGKWRPSPGDFPCIPDRYRAOR 120
 DB 59 HFTRYVTDGCRSAKPVTETVCSGCCGPARRLLPVAIGRGKWRPSPGDFPCIPDRYRAOR 118
 QY 121 VOLLCPGGEAPRAKRVLVASCKCKRLTRFNOSSELDGFTGEARPOKGRKPRRARSAR 180
 DB 119 VOLLCPGGEAPRAKRVLVASCKCKRLTRFNOSSELDGFTGEARPOKGRKPRRARSAR 175
 QY 181 ANQAELENAY 190
 DB 176 ANQAELENAY 185

RESULT 4

US-09-867-274-6
 ; Sequence 6, Application US/09867274
 ; Patent No. US20020106650A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Paszty, Christopher
 ; APPLICANT: Gao, Yongming
 ; TITLE OF INVENTION: Cysteine Knot Polypeptides: Cloaked-2 Molecules and Uses Thereof
 ; FILE REFERENCE: 01017/37428
 ; CURRENT APPLICATION NUMBER: US/09/867,274
 ; PRIOR FILING DATE: 2001-05-29
 ; PRIOR APPLICATION NUMBER: US 60/208,550
 ; PRIOR FILING DATE: 2000-06-01
 ; PRIOR APPLICATION NUMBER: US 60/223,542
 ; PRIOR FILING DATE: 2000-08-04
 ; NUMBER OF SEQ ID NOS: 25
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 6
 ; LENGTH: 208
 ; TYPE: PRT
 ; ORGANISM: Mus musculus
 US-09-867-274-6

Query Match 88.9%; Score 932.5; DB 10; Length 208;
 Best Local Similarity 88.4%; Pred. No. 1.7e-84;
 Matches 168; Conservative 10; Mismatches 7; Indels 5; Gaps 2;

QY 1 OGQWAFKNDATETIIPELGEYEPPEPELENNKTNRAENGRRPHHPETKDVSEYSCREL 60
 DB 24 OGQWAFKNDATETIIPELGEYEPPEPELENNKTNRAENGRRPHHPETKDVSEYSCREL 81
 QY 61 HFTRYVTDGCRSAKPVTETVCSGCCGPARRLLPVAIGRGKWRPSPGDFPCIPDRYRAOR 120
 DB 82 HFTRYVTDGCRSAKPVTETVCSGCCGPARRLLPVAIGRGKWRPSPGDFPCIPDRYRAOR 141
 QY 121 VOLLCPGGEAPRAKRVLVASCKCKRLTRFNOSSELDGFTGEARPOKGRKPRRARSAR 180
 DB 142 VOLLCPGGEAPRAKRVLVASCKCKRLTRFNOSSELDGFTGEARPOKGRKPRRARSAR 198
 QY 181 ANQAELENAY 190
 DB 199 ANQAELENAY 208

RESULT 5

US-09-864-761-47109
 ; Sequence 47109, Application US/09864761
 ; Patent No. US20020048763A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Penn, Sharron G.
 ; APPLICANT: Rank, David R.
 ; APPLICANT: Hanzel, David K.
 ; APPLICANT: Chen, Wensheng
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
 ; FILE REFERENCE: Aecm1ca-X-1


```

; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 47109
; LENGTH: 139
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC003098.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.62
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.6
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.49
; OTHER INFORMATION: SWISSPROT HIT: P45646, EVALU6 4.70e-01
; US-09-864-761-47109

Query Match      72.3%; Score 758; DB 10; Length 139;
Best Local Similarity 100.0%; Pred. No. 1.7e-67;
Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 52 VSEYSGRELFHTYVTDGPGRSAPVTELVCSGCGPARLLPNAIGRGKWRSPGDFRC 111
DB 1 VSEYSGRELFHTYVTDGPGRSAPVTELVCSGCGPARLLPNAIGRGKWRSPGDFRC 60

QY 112 IPDPRYARQVOLLCPGGEAPRAKRVLVASCKKRLTRFNOSLKDFTGEARPOKGRK 171
DB 61 IPDPRYARQVOLLCPGGEAPRAKRVLVASCKKRLTRFNOSLKDFTGEARPOKGRK 120

QY 172 PRPRASAKANOAELENAY 190
DB 121 PRPRASAKANOAELENAY 139
```

```

RESULT 6
US-09-866-050A-159
; Sequence 159, Application US/09866050A
; Publication No. US20030040471A1
```

```

; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Ornust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; FILE REFERENCE: 11000.1011c4U
; CURRENT APPLICATION NUMBER: US/09/866,050A
; CURRENT FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 725
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 159
; LENGTH: 206
; TYPE: PRT
; ORGANISM: Mouse
; US-09-866-050A-159

Query Match      32.8%; Score 344; DB 9; Length 206;
Best Local Similarity 42.7%; Pred. No. 1.9e-26;
Matches 79; Conservative 30; Mismatches 62; Indels 14; Gaps 8;

QY 5 AFKNDATETIPELGEYPP-PPELENNKTMRAENGRRPHHPETKDV---EYSCREL 60
DB 23 AFKNDATETI--YSHVVPVPAHPSSNSTLNQARNGR--HFSSTGLDRNSRVYGCREL 78

QY 61 HTRVYVTDGPGRSAPVTELVCSGCGPARLLPNAIGRG---KWM-RPSGDFCIPRY 116
DB 79 RSTKYSIQGCTSLSPLELVACGCLPLVLPWMTGGYGTCTKWSRSGEMRCVNDKT 138

QY 117 RAQRVOLLCPGGEAPRAKRVLVASCKKRLTRFNOSLKDFTGEARPOKGRKPRRA 176
DB 139 RFRQRLQCGQGS-STRTYKTIYVTAACKCKRTTRQNESNHFEVSAPAKPAQHRRERGA 197

QY 177-RSAR 180
DB 198 SKSSK 202
```

```

RESULT 7
US-09-866-050A-286
; Sequence 286, Application US/09866050A
; Publication No. US20030040471A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Ornust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; FILE REFERENCE: 11000.1011c4U
; CURRENT APPLICATION NUMBER: US/09/866,050A
; CURRENT FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 725
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 286
; LENGTH: 206
; TYPE: PRT
; ORGANISM: Mouse
; US-09-866-050A-286

Query Match      32.8%; Score 344; DB 9; Length 206;
Best Local Similarity 42.7%; Pred. No. 1.9e-26;
Matches 79; Conservative 30; Mismatches 62; Indels 14; Gaps 8;

QY 5 AFKNDATETIPELGEYPP-PPELENNKTMRAENGRRPHHPETKDV---EYSCREL 60
DB 23 AFKNDATETI--YSHVVPVPAHPSSNSTLNQARNGR--HFSSTGLDRNSRVYGCREL 78
```


US-09-867-274-25
 ; Sequence 25, Application US/09867274
 ; Patent No. US20020106650A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Paszy, Christopher
 ; APPLICANT: Gao, Yongming
 ; TITLE OF INVENTION: Cysteine Knot Polypeptides: Cloaked-2 Molecules and Uses Thereof
 ; FILE REFERENCE: 01017/37428
 ; CURRENT APPLICATION NUMBER: US/09/867,274
 ; CURRENT FILING DATE: 2001-05-29
 ; PRIOR APPLICATION NUMBER: US 60/208,550
 ; PRIOR FILING DATE: 2000-06-01
 ; PRIOR APPLICATION NUMBER: US 60/223,542
 ; PRIOR FILING DATE: 2000-08-04
 ; NUMBER OF SEQ ID NOS: 25
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 25
 ; LENGTH: 183
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-867-274-25

Query Match 32.1%; Score 337; DB 10; Length 183;
 Best Local Similarity 42.4%; Pred. No. 7, 9e-26;
 Matches 78; Conservative 30; Mismatches 62; Indels 14; Gaps 8;

QY 6 FKNDATETIPELGEYPER-PELENNKTKMRAENGRRPHHPETKDVSEYSCRELIHT-RR 61
 DB 1 FKNDATETIPEL-GEYPER-PELENNKTKMRAENGRRPHHPETKDVSEYSCRELIHT-RR 56
 QY 62 FTFTYTDGPRSAKPVTELVCSGCCPARLLPNAIGRG--KMW-RPSGDFRCIPDRYR 117
 DB 57 STYKIDSGQCTSTSLPELVYACAGECLPVLPLWIGGYCTKTWSRRSSQEWRCVWDKTR 116
 QY 118 AGFVOLLCPGSEAPRAKRVLVASCCKRLTRPHNOSELKDFTEARPOKGRKPRRA- 176
 DB 117 TORIQLOCCQD-STRTYKIVTVYACKCKRYTRQHNESHNFESMSPAKPVQHRERKRAS 175
 QY 177 RSKA 180
 DB 176 KSKK 179

RESULT 11
 US-09-814-777A-126
 ; Sequence 126, Application US/09814777A
 ; Patent No. US20020142415A1
 ; GENERAL INFORMATION:
 ; APPLICANT: KOOPMAN, Peter Anthony
 ; APPLICANT: MUSCAT, George Eugene Orlando
 ; TITLE OF INVENTION: NOVEL POLYPEPTIDES AND POLYNUCLEOTIDES AND METHODS OF USING THEM
 ; FILE REFERENCE: 21415-0003
 ; CURRENT APPLICATION NUMBER: US/09/814,777A
 ; CURRENT FILING DATE: 2001-03-23
 ; PRIOR APPLICATION NUMBER: AU P06457
 ; PRIOR FILING DATE: 2000-03-24
 ; NUMBER OF SEQ ID NOS: 128
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 126
 ; LENGTH: 337
 ; TYPE: PRT
 ; ORGANISM: Human
 US-09-814-777A-126

Query Match 9.3%; Score 98; DB 10; Length 337;
 Best Local Similarity 25.8%; Pred. No. 0.063;
 Matches 48; Conservative 15; Mismatches 57; Indels 66; Gaps 10;

QY 13 IIPLEGYPPPEPLENNKTKMRAENGRRPHHPETKDVSEYSCRELIHTRYVTDG--- 69
 DB 129 LLEPLAP-PPPPPE-----DPPAASGARAFRELPLPLGAFDGLGL 168
 QY 70 PCRSKAEVTELVCSGCCPARLL-PNAIGRGKMWRRPSGDFRCIPDRYRA-----QRYVOLL 124

DB 169 FTPEPSPLDGL-----EPGEAFAFPPEPAPRTARWRPSAP-----PLAHRVYVAGPRRLIG 219
 QY 125 CPFG-----EAPRAKRVLVAS-----CKKRLTRFHNSLEKDGTEAARPOKGRK 171
 DB 220 SPGGAGODRAPARSLACTTAPMAPRPARPARCRR-----RPRPRWRA 263
 QY 172 PPRPAR 177
 DB 264 PSPGAR 269

RESULT 12
 US-10-044-716-10
 ; Sequence 10, Application US/10044716
 ; Patent No. US20020159986A1
 ; GENERAL INFORMATION:
 ; APPLICANT: LANGENFELD, John
 ; TITLE OF INVENTION: BONE MORPHOGENETIC PROTEIN-2 IN THE TREATMENT AND DIAGNOSIS OF CM
 ; FILE REFERENCE: 270/070US
 ; CURRENT APPLICATION NUMBER: US/10/044,716
 ; CURRENT FILING DATE: 2002-01-11
 ; PRIOR APPLICATION NUMBER: US60/261,252
 ; PRIOR FILING DATE: 2001-01-12
 ; NUMBER OF SEQ ID NOS: 16
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 10
 ; LENGTH: 184
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-044-716-10

Query Match 8.5%; Score 89.5; DB 9; Length 184;
 Best Local Similarity 24.1%; Pred. No. 0.021;
 Matches 38; Conservative 22; Mismatches 57; Indels 41; Gaps 9;

QY 18 GEYPPPELENNK-----TNNRAENGRRPHHPETKDVSEYSCRELIHT-RR 64
 DB 31 GAIP-PPKAGHNDSEQTPQPSRNRGRGGGRGTAMPGE--EVLESQDALHVTERR 87
 QY 65 YTDGPCR-----SAKPVTELVCSGCCPARLLPNAIGRGKMWRRPSGDFR- 110
 DB 88 YLRDWCCKTQPLKQTHIEGCGNSRTIINRCYGCQCN-SFYIPIRHL-----RKEGSGFS 140
 QY 111 ---CIPDRYRAQRYVOLLCPGSEAP-RARKRVLVASCKC 144
 DB 141 CSFCCKKFTTMMVTTLNCELOPFTKGRVTRVKQCR 178

RESULT 13
 US-10-159-749-2
 ; Sequence 2, Application US/10159749
 ; Publication No. US20020192219A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Harland, Richard
 ; APPLICANT: Hsu, David
 ; TITLE OF INVENTION: Morphogenic Proteins
 ; NUMBER OF SEQUENCES: 13
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESS: SCIENCE & TECHNOLOGY LAW GROUP
 ; STREET: 75 DENISE DRIVE
 ; CITY: HILLSBOROUGH
 ; STATE: CALIFORNIA
 ; COUNTRY: USA
 ; ZIP: 94010
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/10/159,749
 ; FILING DATE: 29-May-2002

CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/040,229
FILING DATE: 13-Mar-1998
ATTORNEY/AGENT INFORMATION:
NAME: OSMAN, RICHARD A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: B97-020-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 343-4341
TELEFAX: (650) 343-4342
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 184 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-159-749-2

Query Match 8.5%; Score 89.5; DB 9; Length 184;
Best Local Similarity 24.1%; Pred. No. 0.21;
Matches 38; Conservative 22; Mismatches 57; Indels 41; Gaps 9;

QY 18 GEYEPPELENNK-----TNRRAENGRRPHHPETKDVSEYSCRELIHT--R 64
DB 31 GAIP-PPKAGHNDSEQTSPQSPGSRNRGRGTAMPGE--EVLSSQELHITERK 87

QY 65 YVTGDPGR-----SAKPVTELVSGGCGPARLLPNAIGKMMRPSGDPFR- 110
DB 88 YLRKDWCKTQPLKQTHIEGNSRTIIRFCYGCN-SFYIPRH-----RKEGSGFQS 140

QY 111 --CIPRYRAQRVQLCPGGEAP-RARKVLVASCRC 144
DB 141 CSFCKPKKFTTMTVTLNCPELQPRKKKRVTRVACRC 178

RESULT 14
US-10-159-749-6
Sequence 6, Application US/10159749
Publication No. US20020192219A1
GENERAL INFORMATION:
APPLICANT: Harland, Richard
Hsu, David
TITLE OF INVENTION: Morphogenic Proteins
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 75 DENISE DRIVE
CITY: HILLSBOROUGH
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94010
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/159,749
FILING DATE: 29-May-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/040,229
FILING DATE: 13-Mar-1998
ATTORNEY/AGENT INFORMATION:
NAME: OSMAN, RICHARD A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: B97-020-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 343-4341
TELEFAX: (650) 343-4342
INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:
LENGTH: 184 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-10-159-749-6

Query Match 8.3%; Score 87.5; DB 9; Length 184;
Best Local Similarity 22.4%; Pred. No. 0.32;
Matches 34; Conservative 24; Mismatches 55; Indels 39; Gaps 8;

QY 23 PPELE-----NNKTNRAENGRRPHHPETKDVSEYSCRELIHT--RYTDP 70
DB 36 PPDDKQPNDSHQWQOQSGSRHRERKGTSMPE--EVLSSQELHITERK 93

QY 71 CR-----SAKPVTELVSGGCGPARLLPNAIGKMMRPSGDPFR-----CIP 113
DB 94 CKTQPLKQTHIEGNSRTIIRFCYGCN-SFYIPRH-----RKEGSGFQS 146

QY 114 DRYRAQRVQLCPGGEAPRARK-VRLVASCRC 144
DB 147 KFTTMTVTLNCPELQPRKKKRVTRVACRC 178

RESULT 15
US-10-159-749-4
Sequence 4, Application US/10159749
Publication No. US20020192219A1
GENERAL INFORMATION:
APPLICANT: Harland, Richard
Hsu, David
TITLE OF INVENTION: Morphogenic Proteins
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 75 DENISE DRIVE
CITY: HILLSBOROUGH
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94010
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/159,749
FILING DATE: 29-May-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/040,229
FILING DATE: 13-Mar-1998
ATTORNEY/AGENT INFORMATION:
NAME: OSMAN, RICHARD A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: B97-020-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 343-4341
TELEFAX: (650) 343-4342
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 182 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-159-749-4

Query Match 8.2%; Score 86.5; DB 9; Length 182;
Best Local Similarity 22.8%; Pred. No. 0.4;
Matches 34; Conservative 24; Mismatches 64; Indels 27; Gaps 7;

```

OY 18 GEYEPPEPELENNKTNRAENG---GRPPHHPFETKDVSEYSCRELAHT--RYVTDGPC 71
Db 33 GAIPPDKGOPNDSEOGQAOPEGDVRVKGKGQALAEVLESQALHTTERKYLKRDWC 92
OY 72 R-----SAKPYTELVCSGQCGPARLLPNAIGR--GKWTREPSPDPFCIPDRY 116
Db 93 KTOPLKQTIHEDGCSNRTIINRPGYQCN-SFYIPRHIREEGSFQCS---FCRPKKF 147
OY 117 RAORVOLCPGGEAP-RARKVRLVASCRC 144
Db 148 TTMVVTIINCPELOPFTKKKRTVVKQCRC 176

```

Search completed: March 28, 2003, 14:20:45
 Job time : 14.16 secs

GenCore version 5.1.4 p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 28, 2003, 14:16:57 ; Search time 33.44 Seconds
(without alignments)
757.106 Million cell updates/sec

Title: US-09-867-274-2

Perfect score: 190
Sequence: 1 GCMQAFKNDATETIPELGEV.....KPRPRASAKNQAELENAY 190

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 908470 seqs, 133250620 residues

Word size : 0

Total number of hits satisfying chosen parameters: 903470

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : A_Geneseq_101002.*

1: /SID2/gcgdata/geneseq-emb1/AA1980.DAT:*
2: /SID2/gcgdata/geneseq-emb1/AA1981.DAT:*
3: /SID2/gcgdata/geneseq-emb1/AA1982.DAT:*
4: /SID2/gcgdata/geneseq-emb1/AA1983.DAT:*
5: /SID2/gcgdata/geneseq-emb1/AA1984.DAT:*
6: /SID2/gcgdata/geneseq-emb1/AA1985.DAT:*
7: /SID2/gcgdata/geneseq-emb1/AA1986.DAT:*
8: /SID2/gcgdata/geneseq-emb1/AA1987.DAT:*
9: /SID2/gcgdata/geneseq-emb1/AA1988.DAT:*
10: /SID2/gcgdata/geneseq-emb1/AA1989.DAT:*
11: /SID2/gcgdata/geneseq-emb1/AA1990.DAT:*
12: /SID2/gcgdata/geneseq-emb1/AA1991.DAT:*
13: /SID2/gcgdata/geneseq-emb1/AA1992.DAT:*
14: /SID2/gcgdata/geneseq-emb1/AA1993.DAT:*
15: /SID2/gcgdata/geneseq-emb1/AA1994.DAT:*
16: /SID2/gcgdata/geneseq-emb1/AA1995.DAT:*
17: /SID2/gcgdata/geneseq-emb1/AA1996.DAT:*
18: /SID2/gcgdata/geneseq-emb1/AA1997.DAT:*
19: /SID2/gcgdata/geneseq-emb1/AA1998.DAT:*
20: /SID2/gcgdata/geneseq-emb1/AA1999.DAT:*
21: /SID2/gcgdata/geneseq-emb1/AA2000.DAT:*
22: /SID2/gcgdata/geneseq-emb1/AA2001.DAT:*
23: /SID2/gcgdata/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	190	100.0	190	23	Human cloaked-2 po
2	190	100.0	213	21	Human DAN/Cerberus
3	190	100.0	213	21	Human TGF-beta bin
4	190	100.0	213	22	Human secreted pro
5	190	100.0	213	23	Human pro peptide
6	190	100.0	213	23	Human cloaked-2 po
7	190	100.0	213	23	Human osteocalcin p
8	175	92.1	213	21	Human TGF-beta bin
9	175	92.1	213	21	Human TGF-beta bin
10	139	73.2	139	22	Human bone marrow

11	139	73.2	139	23	ABG42911	Human peptide enco
12	139	73.2	367	21	AA826105	Human DAN/Cerberus
13	128	67.4	213	21	AA96431	Verret TGF-beta bi
14	76	40.0	176	21	AA96434	Bovine TGF-beta bi
15	33	17.4	213	21	AA96433	Rat TGF-beta bindi
16	32	16.8	188	23	AB807208	Mouse cloaked-2 po
17	32	16.8	211	21	AA96432	Mouse TGF-beta bi
18	32	16.8	211	23	AB807210	Mouse cloaked-2 po
19	26	13.7	50	20	AA912009	Human 5' EST seque
20	26	13.7	116	18	AA97654	Secreted protein A
21	9	4.7	116	18	AA944090	Human secreted pro
22	9	4.7	206	19	AA909408	Human small CCN-11
23	9	4.7	206	19	AA958704	Human small CCN-11
24	9	4.7	206	21	AA95711	Human adult retina
25	9	4.7	206	21	AA910233	Human adult retina
26	9	4.7	206	21	AA975981	Murine skin cell p
27	9	4.7	206	21	AA976031	Murine skin cell p
28	9	4.7	206	22	AA955920	Skin cell protein,
29	9	4.7	206	22	AA955970	Human angio genesis
30	9	4.7	206	23	AB95457	Human small CCN-11
31	9	4.7	206	23	AB809200	SCGF CNV family pr
32	9	4.7	206	23	AB809210	Human PRO532 prote
33	9	4.7	206	23	AB844851	Murine protein iso
34	9	4.7	206	23	AB872120	Murine protein iso
35	9	4.7	206	23	AB872130	Human breast tumou
36	9	4.7	206	23	AA982653	Human breast tumou
37	8	4.2	183	23	AB807213	Human cloaked-1 po
38	8	4.2	428	20	AA952704	Human ELK-1 protei
39	8	4.2	428	20	AA952824	Human ELK-1. Homo
40	7	3.7	93	22	AA003012	Human polypeptide
41	7	3.7	182	23	AA982655	Human breast tumou
42	7	3.7	186	19	AA965018	Forsthia dirigent
43	7	3.7	186	22	AA965531	Forsthia dirigent
44	7	3.7	186	22	AA965531	Protein of divergen
45	7	3.7	205	22	AB930347	Peptide #2398 enco

ALIGNMENTS

RESULT 1	AB807207	AB807207 standard; Protein, 190 AA.
ID	AB807207	
XX	AC	AB807207;
XX	DT	26-MAR-2002 (first entry)
XX	DE	Human cloaked-2 polypeptide mature protein sequence.
XX	KW	Cloaked-2; cysteine knot motif; nephrotoxic; cardiatic; immunomodulator;
KW	hepatocarcinoma; antiinflammatory; antithyroid; cytoskeletal; neuroprotective;	
KW	antidiabetic; hypotensive; antidiabetic; antidiabetic; antidiabetic; muscular;	
KW	antidiabetic; anorectic; gene therapy; cell therapy; antisense therapy;	
human.		
XX	OS	Homo sapiens.
XX	FT	Key
XX	FT	Misc-difference 1.16
XX	FT	Location/Qualifiers
XX	FT	Figure"
XX	FT	"this region has been repeated twice in the
XX	FT	sequence provided in the sequence listing but has
XX	FT	been indicated correctly in the sequence in the
XX	FT	Figure"
XX	PN	WO200192308-A2.
XX	PD	06-DEC-2001.
XX	PF	29-MAY-2001; 2001WO-US17478.
XX	PR	01-JUN-2000; 2000US-208550P.
XX	PR	04-AUG-2000; 2000US-223542P.

XX (AMGE-) AMGEN INC.
 PA Paszty CJ, Gao Y;
 PI WPI; 2002-114325/15.
 DR N-PSDB; ABA94293.
 XX
 PT New human and mouse cysteine-knot polypeptide designated as Cloaked-2,
 PT for treating or preventing kidney, heart (e.g. myocardial infarction)
 PT or liver (e.g. hepatitis) diseases
 XX
 PS Claim 13; Fig 1; 170pp; English.
 XX
 CC The invention relates to polypeptides comprising a cysteine knot motif
 CC and designated as Cloaked-2, derived from human and mouse. The cloaked-2
 CC polypeptides can be expressed by standard recombinant methodology. The
 CC cloaked-2 polynucleotides are useful in gene therapy and antisense
 CC therapy. The cloaked-2 polypeptides and polynucleotides are useful for
 CC treating, preventing, ameliorating or detecting diseases and disorders of
 CC the kidney (e.g. anemia, hypertension or low blood pressure), heart (e.g.
 CC cardiac hypertrophy, congestive heart failure, myocardial infarction,
 CC arrhythmias, atherosclerosis, hypertension or low blood pressure),
 CC skeletal muscle (e.g. muscular dystrophy or cachexia), placenta (e.g.
 CC congenital abnormalities or miscarriage), liver (e.g. hepatitis or
 CC cirrhosis), pancreas (e.g. diabetes or pancreatitis), thyroid (e.g.
 CC Graves' disease or myxedema) or adrenal cortex (e.g. Cushing's disease
 CC or Addison's disease), homeostasis or metabolic diseases (e.g. obesity,
 CC cancer or myopathies), infections, or autoimmune diseases. Selective
 CC binding agents may be used to modulate the biological activities of
 CC Cloaked-2 polypeptides or to detect Cloaked-2 polypeptide levels in a
 CC sample. Transgenic non-human animals are useful for drug candidate
 CC screening. The present sequence represents the human cloaked-2
 CC mature polypeptide sequence.

XX Sequence 190 AA;

Query Match 100.0%; Score 190; DB 23; Length 190;

Best Local Similarity 100.0%; Pred. No. 1.9e-190; Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGWAQFNKDATEIIPELGEYEPPELENKTKMRAENGGRPPHHPETKDVSEYSQREL 60
 DB 1 QGWAQFNKDATEIIPELGEYEPPELENKTKMRAENGGRPPHHPETKDVSEYSQREL 60
 QY 61 HFTRYVTGPGCSAKPVTIELVCSGCGPARLLPNAIGKGMWRSGDFRCIPDRYRAQR 120
 DB 61 HFTRYVTGPGCSAKPVTIELVCSGCGPARLLPNAIGKGMWRSGDFRCIPDRYRAQR 120
 QY 121 VOLLCPGSEAPRARKVRLVASCKCKRLTRFNOSSELDQFTGARPPQKGRKPRPARSAK 180
 DB 121 VOLLCPGSEAPRARKVRLVASCKCKRLTRFNOSSELDQFTGARPPQKGRKPRPARSAK 180
 QY 181 ANQAELENAY 190
 DB 181 ANQAELENAY 190

RESULT 2
 AAB26106
 ID AAB26106 standard; Protein; 213 AA.

XX AAB26106;

XX 15-JAN-2001 (first entry)

XX Human DAN/Cerberus-related protein 6 (hDCR6) (exons 1 and 4).

XX Human; DNA/Cerberus-related protein 6; hDCR6; morphogenic protein;
 KW antagonist; BMP; cell growth; cell differentiation; bone formation;
 KM gene therapy.
 XX
 OS Homo sapiens.

XX WO20005193-A2.
 PN 21-SEP-2000.
 PD 02-MAR-2000; 2000WO-US05537.
 XX 12-MAR-1999; 99US-0124118.
 PR (REG-) REGENERON PHARM INC.
 XX
 PA Economides AN;
 PI WPI; 2000-638179/61.
 DR N-PSDB; AAA94051.
 XX
 PT Novel isolated, human DNA/Cerberus related protein 6 which include
 PT natural homologue, and polypeptides comprising DCR6 domain and nucleic
 PT acids encoding the proteins which are useful as probes and primers -

XX Claim 8; Fig 3; 40pp; English.

XX The present sequence comprises the amino acid sequence encoded by exons 1
 CC and 4 of the human DAN/Cerberus-related protein 6 (hDCR6) coding
 CC sequence. The coding sequence was isolated from a human kidney cDNA
 CC library containing exons 1 and 4 of the sequence. hDCR6 is closely
 CC related to the DAN and DCR5 proteins, both of which act as antagonists of
 CC morphogenic proteins such as BMP. It is possible that the hDCR6 gene and
 CC protein can be used as immunogens, modulators of cell function, growth
 CC and differentiation, to reduce undesirable bone formation, to identify
 CC DCR6 binding agents, in diagnosis, and in gene therapy.

XX Sequence 213 AA;

Query Match 100.0%; Score 190; DB 21; Length 213;

Best Local Similarity 100.0%; Pred. No. 2.1e-190; Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGWAQFNKDATEIIPELGEYEPPELENKTKMRAENGGRPPHHPETKDVSEYSQREL 60
 DB 24 QGWAQFNKDATEIIPELGEYEPPELENKTKMRAENGGRPPHHPETKDVSEYSQREL 83
 QY 61 HFTRYVTGPGCSAKPVTIELVCSGCGPARLLPNAIGKGMWRSGDFRCIPDRYRAQR 120
 DB 84 HFTRYVTGPGCSAKPVTIELVCSGCGPARLLPNAIGKGMWRSGDFRCIPDRYRAQR 143
 QY 121 VOLLCPGSEAPRARKVRLVASCKCKRLTRFNOSSELDQFTGARPPQKGRKPRPARSAK 180
 DB 144 VOLLCPGSEAPRARKVRLVASCKCKRLTRFNOSSELDQFTGARPPQKGRKPRPARSAK 203
 QY 181 ANQAELENAY 190
 DB 204 ANQAELENAY 213

RESULT 3
 AAY96429
 ID AAY96429 standard; Protein; 213 AA.

XX AAY96429;

XX 12-SEP-2000 (first entry)

XX Human TGF-beta binding protein (BEER).

XX osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
 KW BEER; gene therapy; antisense therapy; fracture; bone mineralization.

XX Homo sapiens.
 OS WO200032773-A1.
 PN 08-JUN-2000.
 XX

XX 24-NOV-1999; 99WO-US27990.
 PF 27-NOV-1998; 98US-0110283.
 PR (DARM-) DARWIN DISCOVERY LTD.
 PA
 XX Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepfer BW;
 PI Van Ness J, Winkler DG;
 DR N-PSDB; AAA29055.
 DR WPI; 2000-412321/35.
 XX Nucleic acids (1) encoding a transforming growth factor beta binding
 PT protein, useful for identifying agents for treating osteopenia,
 PT osteoporosis and fractures
 XX
 PS Claim 2; Page 116; 162pp; English.
 CC This shows the human transforming growth factor-beta (TGF-beta)
 CC binding protein designated hBER. The cDNA and protein may be used for
 CC prevention, treatment and diagnosis of diseases associated with
 CC inappropriate BBR expression. For example, they may be used to treat
 CC disorders associated with decreased TGF-beta BP expression. The cDNA or
 CC vectors may be administered to treat diseases by rectifying mutations or
 CC deletions in a patient's genome that affect the activity of BBR by
 CC expressing inactive proteins or to supplement the patient's own production
 CC of BBR polypeptides. The nucleic acids may be used for recombinant
 CC production of BBR, gene therapy, antisense therapy, as probes for
 CC diagnostic assays and for functional studies. BBR may be used to raise
 CC antibodies and for identification of BBR modulators. BBR antagonists
 CC may be used to increase bone mineral content for the treatment of
 CC disorders such as osteopenia, osteoporosis, fractures and other
 CC disorders associated with low mineral content.
 XX
 SQ Sequence 213 AA;
 Query Match 100.0%; Score 190; DB 21; Length 213;
 Best Local Similarity 100.0%; Pred. No. 2,1e-190;
 Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 OGMQAFKNDATETIIPELGEYBPPELENNKTNRAENGSRPHHPFTKDVSEYSCREL 60
 DB 24 OGMQAFKNDATETIIPELGEYBPPELENNKTNRAENGSRPHHPFTKDVSEYSCREL 83
 QY 61 HFTFRVYTDGPCRSAPVTELVCSGCGPARLLPNAIGRGKMWRRSPGDPFRCTPDRYRAQR 120
 DB 84 HFTFRVYTDGPCRSAPVTELVCSGCGPARLLPNAIGRGKMWRRSPGDPFRCTPDRYRAQR 143
 QY 121 VOLLCGGEAPRARKVRLVASCCKRLTRFNOSLKDFTGEAARPOKGRKPRPARSAK 180
 DB 144 VOLLCGGEAPRARKVRLVASCCKRLTRFNOSLKDFTGEAARPOKGRKPRPARSAK 203
 QY 181 ANQAELENAY 190
 DB 204 ANQAELENAY 213
 RESULT 4
 ID AA97589 standard; Protein; 213 AA.
 XX AA97589;
 AC
 XX AA97589;
 DT 05-APR-2001 (first entry)
 XX
 DE Human secreted protein PRO7476.
 XX
 KW Secreted protein; human; PRO protein; neoplastic cell growth; tumour;
 KW proliferation; leukaemia; lymphoid malignancy; inflammatory disorder;
 KW angiogenic disorder; immunologic disorder; PRO7476.
 XX
 OS Homo sapiens.

XX WO200075317-A2.
 PN 14-DEC-2000.
 PD
 XX 15-MAY-2000; 2000WO-US13358.
 PF
 XX 09-JUN-1999; 99US-0138385.
 PR 20-JUL-1999; 99US-0144790.
 PR 03-AUG-1999; 99US-0146843.
 PR 10-AUG-1999; 99US-0148188.
 PR 17-AUG-1999; 99US-0149320.
 PR 17-AUG-1999; 99US-0149327.
 PR 17-AUG-1999; 99US-0149396.
 PR 20-AUG-1999; 99US-0150114.
 PR 31-AUG-1999; 99US-0151700.
 PR 31-AUG-1999; 99US-0151734.
 XX
 PA (GETH) GENENTECH INC.
 PI Borestein DA, Goddard A, Gurney AL, Smith V, Watanabe CX, Wood WI;
 DR WPI; 2001-071075/08.
 DR N-PSDB; AAA91023.
 XX
 PT Antibodies against PRO polypeptides, useful for diagnosing and treating
 PT tumours are associated with gene amplification, neoplastic cell growth
 PT and proliferation in mammals -
 XX
 PS Claim 61; Fig 20; 143pp; English.
 CC This sequence is a human PRO protein of the invention. The PRO
 CC proteins are secreted proteins. Antagonists or antibodies of PRO
 CC polypeptides are useful for diagnosing and treating tumours are
 CC associated with gene amplification, neoplastic cell growth and
 CC proliferation in mammals, and those conditions characterised by
 CC overexpression and/or activation of the amplified genes. Such conditions
 CC include benign or malignant tumours (e.g. renal, liver, kidney, bladder,
 CC breast, gastric, ovarian, colorectal, prostate, pancreatic, lung, vulval,
 CC thyroid, hepatic carcinomas, sarcomas, glioblastomas and various head and
 CC neck tumours); leukaemias and lymphoid malignancies; neuronal, glial,
 CC astrocytal, hypothalamic, and other glandular, macrophageal, epithelial,
 CC stromal and blastocoele disorders; and inflammatory, angiogenic and
 CC immunologic disorders. These may further be used to qualitatively or
 CC quantitatively detect the expression of proteins encoded by the
 CC amplified genes, and in tumour diagnostics or prognostics. The PRO
 CC polypeptide or its antagonist may be used for the preparation of a
 CC medicament in the treatment of a condition, which is responsive to the
 CC PRO polypeptide, its antagonist or anti-PRO antibody.
 XX
 SQ Sequence 213 AA;
 Query Match 100.0%; Score 190; DB 22; Length 213;
 Best Local Similarity 100.0%; Pred. No. 2,1e-190;
 Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 OGMQAFKNDATETIIPELGEYBPPELENNKTNRAENGSRPHHPFTKDVSEYSCREL 60
 DB 24 OGMQAFKNDATETIIPELGEYBPPELENNKTNRAENGSRPHHPFTKDVSEYSCREL 83
 QY 61 HFTFRVYTDGPCRSAPVTELVCSGCGPARLLPNAIGRGKMWRRSPGDPFRCTPDRYRAQR 120
 DB 84 HFTFRVYTDGPCRSAPVTELVCSGCGPARLLPNAIGRGKMWRRSPGDPFRCTPDRYRAQR 143
 QY 121 VOLLCGGEAPRARKVRLVASCCKRLTRFNOSLKDFTGEAARPOKGRKPRPARSAK 180
 DB 144 VOLLCGGEAPRARKVRLVASCCKRLTRFNOSLKDFTGEAARPOKGRKPRPARSAK 203
 QY 181 ANQAELENAY 190
 DB 204 ANQAELENAY 213

RESULT 5
 ABG34061 standard; Protein; 213 AA.
 ID ABG34061, standard; Protein; 213 AA.
 AC ABG34061,
 DT 15-JUL-2002 (first entry)
 DE Human Pro peptide #32.
 KW Human; PRO; secreted protein; transmembrane protein;
 KW genetic disorder; tumour; cancer.
 OS Homo sapiens.
 PN WO200224888-A2.
 PD 28-MAR-2002.
 PF 29-AUG-2001; 2001WO-US27099.
 PR 01-SEP-2000; 2000US-229896P.
 PR 05-SEP-2000; 2000US-230621P.
 PR 22-SEP-2000; 2000US-235147P.
 PR 10-NOV-2000; 2000WO-US30873.
 PR 12-JAN-2001; 2001US-261878P.
 PR 16-JAN-2001; 2001US-261910P.
 PR 16-JAN-2001; 2001US-261939P.
 PR 16-JAN-2001; 2001US-262150P.
 PR 25-JAN-2001; 2001US-264395P.
 PR 02-FEB-2001; 2001US-266421P.
 PR 09-FEB-2001; 2001US-267623P.
 PR 28-FEB-2001; 2001WO-US06520.
 PR 09-MAR-2001; 2001US-274399P.
 PR 03-APR-2001; 2001US-280982P.
 PR 04-APR-2001; 2001US-282129P.
 PR 04-APR-2001; 2001US-282199P.
 PR 09-MAY-2001; 2001US-290589P.
 PR 25-MAY-2001; 2001WO-US17092.
 PR 01-JUN-2001; 2001WO-US17800.
 PR 20-JUN-2001; 2001WO-US19692.
 PR 29-JUN-2001; 2001WO-US21066.
 PR 09-JUL-2001; 2001WO-US21735.
 XX (GENTH) GENENTECH INC.
 PA Baker KP, Baton DL, Filvaroff E, Goddard A, Grimaldi JC,
 PI Gurney AL, Smith V, Stephan J, Watanabe CK, Wood WI, Zhang Z,
 PI Fong S;
 XX WPI; 2002-362426/39.
 DR N-PSDB; ABR69992.
 PT New PRO polypeptides and polynucleotides encoding the polypeptides,
 PT useful in gene therapy, chromosome identification, tissue typing, or
 PT for genetic analysis of individuals with genetic disorders -
 XX
 PS Claim 11, Figure 64; 218pp; English.
 CC This invention relates to the cDNA and protein sequences of novel
 CC secreted and transmembrane polypeptides PRO polypeptides. The
 CC invention also comprises a method for producing the proteins of the
 CC invention by recombinant means and antibodies specific for the protein
 CC of the invention. The antibody may be used for detecting the PRO
 CC proteins of the invention and may be used to modify their activity.
 CC polynucleotides may be used as hybridisation probes for a cDNA library
 CC to isolate the full-length PRO cDNA or to isolate other cDNAs, to
 CC construct hybridisation probes for mapping the gene which encodes that
 CC PRO and for genetic analysis of individuals with genetic disorders, in
 CC assays to identify other proteins or molecules involved in binding
 CC reaction, to generate transgenic animals or knock-out animals which in
 CC turn are useful in the development and screening of therapeutically
 CC useful reagents, for chromosome identification, and tissue typing. The

CC PRO polypeptides are useful in gene therapy, and as molecular weight
 CC markers for protein electrophoresis purposes. The sequences may
 CC also be used to detect overexpression on PRO polypeptides in cancerous
 CC tumours and for screening for differentially expressed genes using
 CC microarray technology. The present sequence represents a human PRO
 CC protein of the invention.
 XX
 SQ Sequence 213 AA;
 Query Match 100.0%; Score 190; DB 23; Length 213;
 Best Local Similarity 100.0%; Pred. No. 2,1e-190;
 Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 OGMQAFKNDATETIIPETIGEYPPPELENNKTKMRAENGGRPHHPETDVSYSREL 60
 DB 24 OGMQAFKNDATETIIPETIGEYPPPELENNKTKMRAENGGRPHHPETDVSYSREL 83
 QY 61 HFTRYVTDPGCRSAKPYTELVCSGGCGPARLLPNAIGRKWRRSGDFRCIPRYAQR 120
 DB 84 HFTRYVTDPGCRSAKPYTELVCSGGCGPARLLPNAIGRKWRRSGDFRCIPRYAQR 143
 QY 121 VOLLCPGGEAPRARKVRLVASCCKRLTRFNQSELDGFTGARPPQGRKPRPARSAK 180
 DB 144 VOLLCPGGEAPRARKVRLVASCCKRLTRFNQSELDGFTGARPPQGRKPRPARSAK 203
 QY 181 ANQAELENAY 190
 DB 204 ANQAELENAY 213
 XX
 RESULT 6
 ABB07209
 ID ABB07209 standard; Protein; 213 AA.
 AC ABB07209;
 DT 26-MAR-2002 (first entry)
 DE Human cloaked-2 polypeptide sequence.
 XX
 KW Cloaked-2; cysteine knot motif; nephrotropic; cardiac; immunomodulator;
 KW hepatotropic; antiinflammatory; antihypertoid; cytostatic; neuroprotective;
 KW antianemic; hypotensive; antiarrhythmic; antiarteriosclerotic; muscular;
 KW antidiabetic; anorectic; gene therapy; cell therapy; antisense therapy;
 KW human.
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FT Peptide 1..23
 FT /note= "signal peptide"
 FT Protein 24..213
 FT /note= "mature protein (ABB07207)"
 PN WO200192308-A2.
 PD 06-DEC-2001.
 PF 29-MAY-2001; 2001WO-US17478.
 PR 01-JUN-2000; 2000US-208550P.
 PR 04-AUG-2000; 2000US-223542P.
 PA (AMGE-) AMGEN INC.
 PI Paszty CJ, Gao Y;
 XX WPI; 2002-114325/15.
 DR N-PSDB; ABA94293.
 XX
 PT New human and mouse cysteine-knot polypeptide designated as Cloaked-2,
 PT for treating or preventing kidney, heart (e.g. myocardial infarction)
 PT or liver (e.g. hepatitis) diseases -

XX Example 1; Fig 1; 170pp; English.

CC The invention relates to polypeptides comprising a cysteine knot motif
CC and designated as Cloaked-2, derived from human and mouse. The cloaked-2
CC polypeptides can be expressed by standard recombinant methodology. The
CC cloaked-2 polynucleotides are useful in gene therapy and antisense
CC therapy. The cloaked-2 polypeptides and polynucleotides are useful for
CC treating, preventing, ameliorating or detecting diseases and disorders of
CC the kidney (e.g. anemia, hypertension or low blood pressure), heart (e.g.
CC cardiac hypertrophy, congestive heart failure, myocardial infarction,
CC arrhythmias, atherosclerosis, hypertension or low blood pressure),
CC skeletal muscle (e.g. muscular dystrophy or cachexia), placenta (e.g.
CC congenital abnormalities or miscarriage), liver (e.g. hepatitis or
CC cirrhosis), pancreas (e.g. diabetes or pancreatitis), thyroid (e.g.
CC Grave's disease or myxedema) or adrenal cortex (e.g. Cushing's disease
CC or Addison's disease), homeostasis or metabolic diseases (e.g. obesity,
CC cancer or myopathies), infections, or autoimmune diseases. Selective
CC binding agents may be used to modulate the biological activities of
CC cloaked-2 polypeptides or to detect cloaked-2 polypeptide levels in a
CC sample. Transgenic non-human animals are useful for drug candidate
CC screening. The present sequence represents the human cloaked-2
CC polypeptide sequence.

XX Sequence 213 AA;

Query Match 100.0%; Score 190; DB 23; Length 213;
Best Local Similarity 100.0%; Pred. No. 2.1e-190;
Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGWAQKNDATETIIPELGEYEPPELENNKTMRAENGGRPHHPFETKDVSEYSCREL 60
DB 24 QGWAQKNDATETIIPELGEYEPPELENNKTMRAENGGRPHHPFETKDVSEYSCREL 83
QY 61 HFTRYVTDGPCRSAPKVTVELVCSGCGPARLLPNAIGRKWMPDSGDPFRCTIPRYRAQR 120
DB 84 HFTRYVTDGPCRSAPKVTVELVCSGCGPARLLPNAIGRKWMPDSGDPFRCTIPRYRAQR 143
QY 121 VOLLCPGGEAPPAKRVRLVASCCKRLTRFNQSELKDPGTEAARPOKGRPRPARSAK 180
DB 144 VOLLCPGGEAPPAKRVRLVASCCKRLTRFNQSELKDPGTEAARPOKGRPRPARSAK 203

QY 181 ANQAELENAY 190
DB 204 ANQAELENAY 213

RESULT 7
AAE17089
ID AAE17089 standard; Protein; 213 AA.

XX AAE17089;

XX 18-APR-2002 (first entry)

XX Human osteolevin protein.

XX Human; osteolevin; osteopathic; cyrostatic; bone formation; osteoporosis;
KW Van Buchem-sclerosteosis disease; sclerosteosis; transgenic animal;
XX Paget's disease; chromosome 17.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Peptide 1..19 /label= signal_peptide

FT Protein 20..213 /label= Mature_osteolevin_protein

FT Misc-difference 10 /note= "During polymorphism wild type Val is
substituted with Ile"

XX WT
XX PN W0200198491-A2.

XX 27-DEC-2001.

PD 15-JUN-2001; 2001WO-EP06795.

XX 19-JUN-2000; 2000EP-0112867.

XX (HOFF) HOFFMANN LA ROCHE & CO AG F.
PA (UYIN-) UNIV INSTELLING ANTWERPEN UTA.

XX Baljans W, Ebeling M, Foerzler D, Patel N, Van Hul W;
PI Vickerly BR;

DR MPI; 2002-139789/18.
N-PSDB; AAD27576, AAD27577.

XX Novel genetic polymorphisms in the Van Buchem-sclerosteosis disease
PT region that are associated with abnormal bone formation useful for
XX diagnosis and assessment of osteoporosis or sclerosteosis in humans
PS Claim 11; Fig 4; 70pp; English.

CC The invention relates to a nucleic acid encoding osteolevin region
CC polymorphisms. The invention also relates to genetic polymorphisms in
CC the Van Buchem-sclerosteosis disease region that are associated with
CC disorders resulting in either net excess bone formation or insufficient
CC bone formation in humans. Osteolevin DNA is useful for screening for
CC osteolevin polymorphisms associated with abnormal bone formation in a
CC subject and for the presence of a heritably linked form of abnormal bone
CC formation in a subject, by determining the presence of a polymorphism in
CC the osteolevin nucleic acid sequence obtained from the subject.
CC Osteolevin protein is useful for treating diseases associated with
CC abnormal bone formation, such as sclerosteosis. Van Buchem's disease and
CC Paget's disease. Nucleic acids which encode genes in the osteolevin
CC region or their modified forms can also be used to generate either
CC transgenic animals or knockout animals which are useful in the screening
CC and development of therapeutically useful reagents. Osteolevin proteins
CC are useful in pharmacological characterisation of novel modulators of the
CC activity of protein and protein complexes. Human osteolevin gene is
CC located on chromosome 17. The present sequence is human osteolevin
CC protein.

XX Sequence 213 AA;

Query Match 100.0%; Score 190; DB 23; Length 213;
Best Local Similarity 100.0%; Pred. No. 2.1e-190;
Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGWAQKNDATETIIPELGEYEPPELENNKTMRAENGGRPHHPFETKDVSEYSCREL 60
DB 24 QGWAQKNDATETIIPELGEYEPPELENNKTMRAENGGRPHHPFETKDVSEYSCREL 83

QY 61 HFTRYVTDGPCRSAPKVTVELVCSGCGPARLLPNAIGRKWMPDSGDPFRCTIPRYRAQR 120

DB 84 HFTRYVTDGPCRSAPKVTVELVCSGCGPARLLPNAIGRKWMPDSGDPFRCTIPRYRAQR 143

QY 121 VOLLCPGGEAPPAKRVRLVASCCKRLTRFNQSELKDPGTEAARPOKGRPRPARSAK 180
DB 144 VOLLCPGGEAPPAKRVRLVASCCKRLTRFNQSELKDPGTEAARPOKGRPRPARSAK 203

QY 181 ANQAELENAY 190
DB 204 ANQAELENAY 213

RESULT 8
AAV96430
ID AAV96430 standard; Protein; 213 AA.

XX AAV96430;

XX 12-SEP-2000 (first entry)

XX

DE Human TGF-beta binding protein (BEER) variant V101.
 XX osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
 KM BEER; variant; V101; gene therapy; antisense therapy; fracture;
 XX chromosome 17q12-21; bone mineralization.
 OS Homo sapiens.
 XX
 XX
 FT Key Location/Qualifiers
 FT Misc-difference 10 /label= V101
 FT /note= "wild type valine has been substituted with
 FT isoleucine"
 XX
 XX MO200032773-A1.
 XX
 XX 08-JUN-2000.
 XX
 XX 24-NOV-1999; 99WO-US27990.
 XX
 XX 27-NOV-1998; 98US-0110283.
 XX
 XX (DARW-) DARWIN DISCOVERY LTD.
 XX
 XX Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepker BW;
 XX Van Ness J, Winkler DG;
 XX WPI; 2000-412321/35.
 XX N-PSDB; AAA29056.
 XX
 XX Nucleic acids (1) encoding a transforming growth factor beta binding
 PT protein, useful for identifying agents for treating osteopenia,
 PT osteoporosis and fractures
 XX
 PS Claim 3; Page 119-120; 162pp; English.
 XX
 XX This shows a variant human transforming growth factor-beta (TGF-beta)
 CC binding protein designated BEER V101, which comprises a substitution of
 CC isoleucine for the wild-type valine at residue 10. The cDNA and protein
 CC may be used for prevention, treatment and diagnosis of diseases
 CC associated with inappropriate BEER expression. For example, they may be
 CC used to treat disorders associated with decreased TGF-beta BP expression.
 CC The cDNA or vectors may be administered to treat diseases by rectifying
 CC mutations or deletions in a patient's genome that affect the activity of
 CC BEER by expressing inactive proteins or to supplement the patients own
 CC production of BEER polypeptides. The nucleic acids may be used for
 CC recombinant production of BEER, gene therapy, antisense therapy, as
 CC probes for diagnostic assays and for functional studies. BEER may be used
 CC to raise antibodies and for identification of BEER modulators. BEER
 CC antagonists may be used to increase bone mineral content for the
 CC treatment of disorders such as osteopenia, osteoporosis, fractures and
 CC other disorders associated with low mineral content.
 CC
 XX
 XX Sequence 213 AA;
 SQ
 Query Match 92.1%; Score 175; DB 21; Length 213;
 Best Local Similarity 100.0%; Pred. No. 1.1e-174;
 Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 16 ELGYPEPPEPELENNKTMNRAENGRRPHHPFETKDSEVSCRELHTRVYTDGPCRSK 75
 DB 39 ELGYPEPPEPELENNKTMNRAENGRRPHHPFETKDSEVSCRELHTRVYTDGPCRSK 98
 QY 76 PYTELVGSGGCGPARLLPNAIGRGKMWRRPSGDPDRCTPDRYRAQVOLLCPGSEAPRARK 135
 DB 99 PYTELVGSGGCGPARLLPNAIGRGKMWRRPSGDPDRCTPDRYRAQVOLLCPGSEAPRARK 158
 QY 136 VRLVASCKCKRLTRFNHOSLKDCTGAARPOKRRKRRPARSKANOAOLENAY 190
 DB 159 VRLVASCKCKRLTRFNHOSLKDCTGAARPOKRRKRRPARSKANOAOLENAY 213
 RESULT 9

AAV96436
 ID AAV96436 standard; Protein; 213 AA.
 XX
 XX AAV96436;
 AC
 XX
 XX 12-SEP-2000 (first entry)
 DT
 XX
 XX Human TGF-beta binding protein (BEER) variant P38R.
 DE
 XX
 XX osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
 KM BEER; variant; P38R; gene therapy; antisense therapy; fracture;
 XX chromosome 17q12-21; bone mineralization.
 XX
 XX
 OS Homo sapiens.
 XX
 XX
 FT Key Location/Qualifiers
 FT Misc-difference 38 /label= P38R
 FT /note= "wild type proline has been substituted with
 FT arginine"
 XX
 XX MO200032773-A1.
 XX
 XX 08-JUN-2000.
 XX
 XX 24-NOV-1999; 99WO-US27990.
 XX
 XX 27-NOV-1998; 98US-0110283.
 XX
 XX (DARW-) DARWIN DISCOVERY LTD.
 XX
 XX Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepker BW;
 XX Van Ness J, Winkler DG;
 XX WPI; 2000-412321/35.
 XX N-PSDB; AAA29062.
 XX
 XX Nucleic acids (1) encoding a transforming growth factor beta binding
 PT protein, useful for identifying agents for treating osteopenia,
 PT osteoporosis and fractures
 XX
 PS Disclosure; Page 121; 162pp; English.
 XX
 XX This shows a variant human transforming growth factor-beta
 CC (TGF-beta) binding protein designated BEER P38R. The encoded protein
 CC comprises a substitution of arginine for the wild-type proline at
 CC residue 38. The cDNA and protein may be used for prevention, treatment
 CC and diagnosis of diseases associated with inappropriate BEER expression.
 CC For example, they may be used to treat disorders associated with
 CC decreased TGF-beta BP expression. The cDNA or vectors may be administered
 CC to treat diseases by rectifying mutations or deletions in a patient's
 CC genome that affect the activity of BEER by expressing inactive proteins
 CC or to supplement the patients own production of BEER polypeptides. The
 CC nucleic acids may be used for recombinant production of BEER, gene
 CC therapy, antisense therapy, as probes for diagnostic assays and for
 CC functional studies. BEER may be used to raise antibodies and for
 CC identification of BEER modulators. BEER antagonists may be used to
 CC increase bone mineral content for the treatment of disorders such as
 CC osteopenia, osteoporosis, fractures and other disorders associated with
 CC low mineral content.
 CC
 XX
 XX Sequence 213 AA;
 SQ
 Query Match 92.1%; Score 175; DB 21; Length 213;
 Best Local Similarity 100.0%; Pred. No. 1.1e-174;
 Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 16 ELGYPEPPEPELENNKTMNRAENGRRPHHPFETKDSEVSCRELHTRVYTDGPCRSK 75
 DB 39 ELGYPEPPEPELENNKTMNRAENGRRPHHPFETKDSEVSCRELHTRVYTDGPCRSK 98
 QY 76 PYTELVGSGGCGPARLLPNAIGRGKMWRRPSGDPDRCTPDRYRAQVOLLCPGSEAPRARK 135

PT Spatially-addressable set of single exon nucleic acid probes, used to
PT measure gene expression in human lung samples -

PS Claim 4; SEQ ID No 19768; 634bp; English.

CC The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human lung comprising single exon nucleic acid probes having one of
CC 12614 nucleic acid sequences mentioned in the specification, or their
CC complements or the 12387 open reading frames derived from the 12614
CC probes. Also included are a microarray comprising the novel set of
CC probes; the novel set of probes which hybridise at high stringency to a
CC nucleic acid expressed in the human lung; measuring gene expression in a
CC sample derived from human lung, comprising (a) contacting the array with
CC a collection of detectably labeled nucleic acids derived from human lung
CC mRNA, and (b) measuring the label detectably bound to each probe of
CC the array; identifying exons in a eukaryotic genome, comprising
CC (a) algorithmically predicting at least one exon from genomic sequences
CC of the eukaryote; and (b) detecting specific hybridisation of detectably
CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC having a fragment identical to the predicted exon, the probe is included
CC in the above mentioned microarray; assigning exons to a single gene,
CC comprising (a) identifying exons from genomic sequence by the method
CC above and (b) measuring the expression of each of the exons in several
CC tissues and/or cell types using hybridisation to a single exon
CC microarrays having a probe with the exon, where a common pattern of
CC expression of the exons in the tissues and/or cell types indicates that
CC the exons should be assigned to a single gene; a peptide comprising one
CC of 12011 sequences, mentioned in the specification, or encoded by the
CC probes/open reading frames (ORF). The probes are used for gene
CC expression analysis, and for identifying exons in a gene, particularly
CC using human lung derived mRNA and for the study of lung diseases
CC such as asthma, lung cancer, chronic obstructive pulmonary disease
CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
CC haemorrhoidosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
CC pulmonary alveolar proteinosis, Karsenger syndrome, fibrocystic
CC and hyaline membrane disease. The present sequence is a single exon
CC probe open reading frame of the invention.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

SO Sequence 422 BP; 53 A; 145 C; 159 G; 65 T; 0 other;

Query Match 55.6%; Score 422; DB 24; Length 422;

Best Local Similarity 100.0%; Pred. No. 3e-186;
Matches 422; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 257 ACCTGTCGAGTACAGTCCGCGAGTGCATTCACCCGCTACGTCGATGGCCGT 316
DB 422 ACCTGTCGAGTACAGTCCGCGAGTGCATTCACCCGCTACGTCGATGGCCGT 363
QY 317 GCCGACGCGCAGACCGGATCAGCAGAGTGTGCTCCGCGCAGTGGACCCGCGCC 376
DB 362 GCCGACGCGCAGACCGGATCAGCAGAGTGTGCTCCGCGCAGTGGACCCGCGCC 303
QY 377 TGTGTCGCAACGCTATGCGCGCGGCAAGTGTGGCGACCTAGTGGACCTCGACT 436
DB 302 TGTGTCGCAACGCTATGCGCGCGGCAAGTGTGGCGACCTAGTGGACCTCGACT 243
QY 437 GCATCCCGACCGCTACCGCGGCGACGCGTGAAGTGTGCTCCGCTGGTGAAGCGCC 496
DB 242 GCATCCCGACCGCTACCGCGGCGACGCGTGAAGTGTGCTCCGCTGGTGAAGCGCC 183
QY 497 CGGCGCGCGCAGAGTGTGCGCTGTGCTGTCGCAAGTGAAGCGCCTTACCCGCTCC 556
DB 182 CGGCGCGCGCAGAGTGTGCGCTGTGCTGTCGCAAGTGAAGCGCCTTACCCGCTCC 123
QY 557 ACAACCAAGTGAAGTCAAGGACTTGGGACCGAGCGCGCTCGGCGCGCAGAGGCGCGGA 616

DB 122 ACAACCAAGTGAAGTCAAGGACTTGGGACCGAGCGCGCTCGGCGCGCAGAGGCGCGGA 63

QY 617 ACCCGCGCGCGCGCGCGCGCGCGGAGCGCCAAACCTACAGCGCGAGTGGAGAACGCTTACT 676

DB 62 ACCCGCGCGCGCGCGCGCGCGCGGAGCGCCAAACCTACAGCGCGAGTGGAGAACGCTTACT 3

QY 677 AG 678

DB 2 AG 1

RESULT 15

AAA94050
ID AAA94050 standard; DNA; 1104 BP.

AAA94050;

15-JAN-2001 (first entry)

Human DAN/Cerberus-related protein 6 (hDCR6) coding sequence #1.

Human, DNA/Cerberus-related protein 6; hDCR6; morphogenic protein;

antagonist; BMP; cell growth; cell differentiation; bone formation;

gene therapy; ss.

Homo sapiens.

Key Location/Qualifiers

CDS 1..1104

FT /tag= a

FT /product= "hDCR6 #1"

PN WO20055193-A2.

PD 21-SEP-2000.

02-MAR-2000; 2000WO-US05537.

PF 12-MAR-1999; 99US-0124118.

PR (REG-) REGENERON PHARM INC.

PA Economides AN;

PI MPI: 2000-638179/61.

DR P-PSDB; AAB26105.

XX Novel isolated, human DNA/Cerberus related protein 6 which include

PT natural homologue, and polypeptides comprising DCR6 domain and nucleic

PS acids encoding the proteins which are useful as probes and primers -

XX Claim 2; Fig 2; 40bp; English.

XX The present sequence comprises the human DAN/Cerberus-related protein 6

CC (hDCR6) coding sequence. It was isolated from a genomic DNA clone

CC following identification using computer-based 'virtual cloning'. hDCR6

CC is closely related to the DAN and DCR5 proteins, both of which act as

CC antagonists of morphogenic proteins such as BMP. It is possible that

CC the hDCR6 gene and protein can be used as immunogens, modulators of cell

CC function, growth and differentiation, to reduce undesirable bone

CC formation, to identify DCR6 binding agents, in diagnosis, and in gene

CC therapy.

Sequence 1104 BP; 226 A; 354 C; 364 G; 160 T; 0 other;

Query Match 55.6%; Score 422; DB 21; Length 1104;

Best Local Similarity 100.0%; Pred. No. 2.7e-186;
Matches 422; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 257 ACCTGTCGAGTACAGTCCGCGAGTGCATTCACCCGCTACGTCGATGGCCGT 316
DB 683 ACCTGTCGAGTACAGTCCGCGAGTGCATTCACCCGCTACGTCGATGGCCGT 742

Qy 317 GCCGACGCGCAAGCCGCTACCGAGCTGGTGTGCTCCGGCCAGTGGGCCCCGGCGGCC 376
|||
Db 743 GCCGACGCGCAAGCCGCTACCGAGCTGGTGTGCTCCGGCCAGTGGGCCCCGGCGGCC 802
|||
Qy 377 TGTGCTCCAAAGCCATCGGCGCGGCAAGTGTGGGCACTTGTGGGCCCGACTTCGGCT 436
|||
Db 803 TGTGCTCCAAAGCCATCGGCGCGGCAAGTGTGGGCACTTGTGGGCCCGACTTCGGCT 862
|||
Qy 437 GCATCCCGGACCGCTACCGGCGCGGCAAGCTGTGTCCCGTGTGGAGGCGC 496
|||
Db 863 GCATCCCGGACCGCTACCGGCGCGGCAAGCTGTGTCCCGTGTGGAGGCGC 922
|||
Qy 497 CGCGCGCGCAAGGTGCGCTGTGGCTCGTGCAGTGCAGCGCTCACCGCTTCC 556
|||
Db 923 CGCGCGCGCAAGGTGCGCTGTGGCTCGTGCAGTGCAGCGCTCACCGCTTCC 982
|||
Qy 557 ACAACCAAGTCGAGCTCAAGGACTTCGGGACCGAGGCCGCTCGGCGCGAGAGGGCCGGA 616
|||
Db 983 ACAACCAAGTCGAGCTCAAGGACTTCGGGACCGAGGCCGCTCGGCGCGAGAGGGCCGGA 1042
|||
Qy 617 AGCGCGCGCCCCCGCGCGAGCGGCCAAAGCCAAACCAAGCCGAGCTGGAGAGAGCCCTACT 676
|||
Db 1043 AGCGCGCGCCCCCGCGCGAGCGGCCAAAGCCAAACCAAGCCGAGCTGGAGAGAGCCCTACT 1102
|||
Qy 677 AG 678
||
Db 1103 AG 1104

Search completed: March 29, 2003, 08:39:17
Job time : 247.266 secs

Db 312 GTACGCGATGGGCGCTGCGCGAGGCGCAAGCCGAGTCAACGAGTGTGTCTCCGCGCAG 371
Qy 361 TGGGCGCCCGCGCGCTGCTGCTGCCCAAGCCATGCGCGCGCAAGTGTGTGCACTAGT 420
Db 372 TGGGCGCCCGCGCGCTGCTGCTGCCCAAGCCATGCGCGCGCAAGTGTGTGCACTAGT 431
Qy 421 GGGCGCGACTTCCGCTGCTGCTGCCCAAGCCATGCGCGCGCAAGTGTGTGCACTAGT 480
Db 432 GGGCGCGACTTCCGCTGCTGCTGCCCAAGCCATGCGCGCGCAAGTGTGTGCACTAGT 491
Qy 481 CCGGCTGTGAGGCGCGCGCGCGCGCAAGTGTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 540
Db 492 CCGGCTGTGAGGCGCGCGCGCGCGCAAGTGTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 551
Qy 541 CGGCTCAACCGCTTCCCAACCAAGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 600
Db 552 CGGCTCAACCGCTTCCCAACCAAGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 611
Qy 601 CCGGAGAAAGGCG 660
Db 612 CCGGAGAAAGGCG 671
Qy 661 CTGGAGAAAGGCG 720
Db 672 CTGGAGAAAGGCG 731
Qy 721 GAACCG 759
Db 732 GAACCG 770

RESULT 2

US-09-449-218D-3
; Sequence 3, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepfer, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 2301
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-449-218D-3

Query Match 93.3%; Score 708; DB 4; Length 2301;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 758; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TACTGGAAGGTGGGCGTCCCTCTGCTGCTGCTACATGAGCTCCGACTGGCCCTGTGT 60
Db 12 TACTGGAAGGTGGGCGTCCCTCTCTGCTGCTGCTACATGAGCTCCGACTGGCCCTGTGT 71
Qy 61 CTGCTGCTGCTGCTGCTGCTACACAGCCTTCCGTGTGTGTGAGGCGCAAGGAGTGTGCAAGCG 120
Db 72 CTGCTGCTGCTGCTGCTGCTACACAGCCTTCCGTGTGTGTGTGAGGCGCTGTGAGGAGTGTGCAAGCG 131
Qy 121 TTCAAGATGTGTGCGAGGAATCATCCCGAGCTCGAGAGTACCCCGAGCTTCCACCG 180
Db 132 TTCAAGATGTGTGCGAGGAATCATCCCGAGCTCGAGAGTACCCCGAGCTTCCACCG 191

Qy 181 GAGCTGGAACAACAAGACCATGAACCGGCGGAGAACGAGGCGGCTCCCGACAC 240
Db 192 GAGCTGGAACAACAAGACCATGAACCGGCGGAGAACGAGGCGGCTCCCGACAC 251
Qy 241 CCCTTTGAGACCAAGAGCTGTCCGAGTACAGCTGTCCGCGAGCTGTGCACTTACCCCGTAC 300
Db 252 CCCTTTGAGACCAAGAGCTGTCCGAGTACAGCTGTCCGCGAGCTGTGCACTTACCCCGTAC 311
Qy 301 GTGACGATGGGCGGTGCGCGAGCGCAAGCGGCTGTGCAAGCTGTGTGCTGTGCTGTGCTGTG 360
Db 312 GTGACGATGGGCGGTGCGCGAGCGCAAGCGGCTGTGCAAGCTGTGTGCTGTGCTGTGCTGTG 371
Qy 361 TGGGCGCGCGCGCGCTGTGCTGCCCAAGCCATGCGCGCGCAAGTGTGTGCACTAGT 420
Db 372 TGGGCGCGCGCGCGCTGTGCTGCCCAAGCCATGCGCGCGCAAGTGTGTGCACTAGT 431
Qy 421 GGGCGCGACTTCCGCTGCTGCTGCCCAAGCCATGCGCGCGCAAGTGTGTGCACTAGT 480
Db 432 GGGCGCGACTTCCGCTGCTGCTGCCCAAGCCATGCGCGCGCAAGTGTGTGCACTAGT 491
Qy 481 CCGGCTGTGAGGCGCGCGCGCGCGCGCAAGTGTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 540
Db 492 CCGGCTGTGAGGCGCGCGCGCGCGCGCAAGTGTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 551
Qy 541 CGGCTCAACCGCTTCCCAACCAAGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 600
Db 552 CGGCTCAACCGCTTCCCAACCAAGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 611
Qy 601 CCGGAGAAAGGCG 660
Db 612 CCGGAGAAAGGCG 671
Qy 661 CTGGAGAAAGGCG 720
Db 672 CTGGAGAAAGGCG 731
Qy 721 GAACCG 759
Db 732 GAACCG 770

RESULT 3

US-09-449-218D-7
; Sequence 7, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepfer, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 2301
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-449-218D-7

Query Match 93.3%; Score 708; DB 4; Length 2301;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 758; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TACTGGAAGGTGGGCGTCCCTCTCTGCTGCTGCTACATGAGCTCCGACTGGCCCTGTGT 60
Db 12 TACTGGAAGGTGGGCGTCCCTCTCTGCTGCTGCTACATGAGCTCCGACTGGCCCTGTGT 71

NUMBER OF SEQ ID NOS: 45
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 9
LENGTH: 642
TYPE: DNA
ORGANISM: Cercopithecus pygerythrus
US-09-449-218D-9

Query Match 14.5%; Score 110; DB 4; Length 642;
Best Local Similarity 99.1%; Pred. No. 1.5e-41;
Matches 210; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 121 TTCAGAAATGATCCCAAGAAATCATCCCGAGCTCGAGAGTACCCGAGCTCCACCG 180
DB 85 TTCAGAAATGATCCCAAGAAATCATCCCGAGCTCGAGAGTACCCGAGCTCCACCG 144
QY 181 GAGCTGGAACAACAAGACATGAAACCGGCGGAGAACGAGAGGGGGCTTCCCAACAC 240
DB 145 GAGCTGGAACAACAAGACATGAAACCGGCGGAGAACGAGAGGGGGCTTCCCAACAC 204
QY 241 CCCTTGAGACCAAAAGACGTGTCGAGTACAGCTGCGGAGCTGACATTACCGGCTAC 300
DB 205 CCCTTGAGACCAAAAGACGTGTCGAGTACAGCTGCGGAGCTGACATTACCGGCTAC 264
QY 301 GTGACCGATGGGCGGTGCGCGACGCGCAAGCC 332
DB 265 GTGACCGATGGGCGGTGCGCGACGCGCAAGCC 296

RESULT 8
US-09-449-218D-15
Sequence 15, Application US/09449218D
Patent No. 6395511

GENERAL INFORMATION:
APPLICANT: Brunkow, Mary E.
APPLICANT: Galas, David J.
APPLICANT: Kovacevich, Brian
APPLICANT: Mulligan, John T.
APPLICANT: Paepert, Bryan W.
APPLICANT: Van Ness, Jeffrey
APPLICANT: Winkler, David G.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
FILE REFERENCE: 240083.508
CURRENT APPLICATION NUMBER: US/09/449,218D
CURRENT FILING DATE: 1999-11-24
NUMBER OF SEQ ID NOS: 45
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 15
LENGTH: 532
TYPE: DNA
ORGANISM: Bos taurus
US-09-449-218D-15

Query Match 10.9%; Score 83; DB 4; Length 532;
Best Local Similarity 99.3%; Pred. No. 3.7e-29;
Matches 133; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 280 GAGCTGACATTCACCCGCTACGATGAGCCGCTGCGCGCCCAAGCCGCTACCC 339
DB 153 GAGCTGACATTCACCCGCTACGATGAGCCGCTGCGCGCCCAAGCCGCTACCC 212
QY 340 GAGCTGATGCTCGGCGCAGTGGCGCGCGCGCTGCTGCCCAAGCAGCATCGGCGCG 399
DB 213 GAGCTGATGCTCGGCGCAGTGGCGCGCGCGCTGCTGCCCAAGCAGCATCGGCGCG 272
QY 400 GGCAGGTGTGGCG 413
DB 273 GGCAGGTGTGGCG 286

RESULT 9
US-09-449-218D-11

Sequence 11, Application US/09449218D
Patent No. 6395511
GENERAL INFORMATION:
APPLICANT: Brunkow, Mary E.
APPLICANT: Galas, David J.
APPLICANT: Kovacevich, Brian
APPLICANT: Mulligan, John T.
APPLICANT: Paepert, Bryan W.
APPLICANT: Van Ness, Jeffrey
APPLICANT: Winkler, David G.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
FILE REFERENCE: 240083.508
CURRENT APPLICATION NUMBER: US/09/449,218D
CURRENT FILING DATE: 1999-11-24
NUMBER OF SEQ ID NOS: 45
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 11
LENGTH: 638
TYPE: DNA
ORGANISM: Mus musculus
US-09-449-218D-11

Query Match 9.1%; Score 69; DB 4; Length 638;
Best Local Similarity 100.0%; Pred. No. 9.7e-23;
Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 517 CTGTGGGCTGTGTGAAGTGAAGGCTTCCACCAACCACTCGAGTCAAG 576
DB 475 CTGTGGGCTGTGTGAAGTGAAGGCTTCCACCAACCACTCGAGTCAAG 534
QY 577 GACTTCGGG 585
DB 535 GACTTCGGG 543

RESULT 10
US-09-449-218D-17
Sequence 17, Application US/09449218D
Patent No. 6395511
GENERAL INFORMATION:
APPLICANT: Brunkow, Mary E.
APPLICANT: Galas, David J.
APPLICANT: Kovacevich, Brian
APPLICANT: Mulligan, John T.
APPLICANT: Paepert, Bryan W.
APPLICANT: Van Ness, Jeffrey
APPLICANT: Winkler, David G.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
FILE REFERENCE: 240083.508
CURRENT APPLICATION NUMBER: US/09/449,218D
CURRENT FILING DATE: 1999-11-24
NUMBER OF SEQ ID NOS: 45
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 17
LENGTH: 35828
TYPE: DNA
ORGANISM: Mus musculus
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)...(35828)
OTHER INFORMATION: n = A,T,C or G
US-09-449-218D-17

Query Match 9.1%; Score 69; DB 4; Length 35828;
Best Local Similarity 100.0%; Pred. No. 7.8e-23;
Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 517 CTGTGGGCTGTGTGAAGTGAAGGCTTCCACCAACCACTCGAGTCAAG 576
DB 21283 CTGTGGGCTGTGTGAAGTGAAGGCTTCCACCAACCACTCGAGTCAAG 21342

QY	577	GACTTCGGG	585
Db	21343	GACTTCGGG	21351

```

RESULT 11
US-09-449-218D-13
; Sequence 13, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepel, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHOD
; OF TITRATION OF BONE MINERALIZATION
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 13
; LENGTH: 674
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-09-449-218D-13

```

Query Match	9.0%;	Score 68;	DB 4;	Length 674;
Best Local Similarity	100.0%;	Pred. No. 2.8e-22;		
Matches 68;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0

```
Oy    517 CTGTGGCCCTCGTGCAAGTGCAGGCGCTTCACCCGTTCCACAACCAGTCGGAGCTCAAG   576
      |||||
Db    513 CTGTGGCCCTCGTGCAAGTGCAGGCGCTTCACCCGTTCCACAACCAGTCGGAGCTCAAG   577
```

Qy	577	GACTTCGG	584
Db	573	GACTTCGG	580

```

RESULT 12
US-09-449-218D-26
; Sequence 26, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepert, Bryan W.
; APPLICANT: Van Nees, Jeffrey
; APPLICANT: Whinkle, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHOD
; OF INVENTION: BONE MINERALIZATION
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218BD
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 26
; LENGTH: 39
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Primer for PCR
US-09-449-218D-26

```

Query Match 3.3%; Score 25; DB 4; Length 39;
Best Local Similarity 100.0%; Pred. No. 0.018;
Matches 25; Conservative 0; Mismatches 0; Indels 0
Gaps 0

Qy 108 GGGGTGCGAGCGCTTCAGAATGAT 132
| | | | | | | | | | | | | |
Db 15 GGGGTGCGAGCGCTTCAGAATGAT 39

```

US-09-449-218D-19
US-09-449-218D-19
: Sequence 19, Application US/09449218D
: Patent No. 6395511
: GENERAL INFORMATION:
: APPLICANT: Brunkow, Mary E.
: APPLICANT: Galas, David J.
: APPLICANT: Kovacevich, Brian
: APPLICANT: Mulligan, John T.
: APPLICANT: Paepert, Bryan W.
: APPLICANT: Van Ness, Jeffrey
: APPLICANT: Winkler, David G.
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
: FILE REFERENCE: 240083.508
: CURRENT APPLICATION NUMBER: US/09/449,218D
: CURRENT FILING DATE: 1999-11-24
: NUMBER OF SEQ ID NOS: 45
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 19
: LENGTH: 21
: TYPE: DNA
: ORGANISM: Artificial Sequence
: FEATURES:
: OTHER INFORMATION: Primer for PCR
US-09-449-218D-19

```

Query Match	2.8%;	Score 21;	DB 4;	Length 21;
Best Local Similarity	100.0%;	Pred. No. 1.3;		
Matches 21;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY 178 CCGAGCTGGAGACAACAAG 198
|||
Db 1 CCGAGCTGGAGACAACAAG 21

```

US-09-449-218D-27/C
Result 14
Sequence 27, Application US/09449218D
Patent No. 6395511
GENERAL INFORMATION:
APPLICANT: Brunkow, Mary E.
APPLICANT: Galas, David J.
APPLICANT: Kovacevich, Brian
APPLICANT: Mulligan, John T.
APPLICANT: Paepert, Bryan W.
APPLICANT: Van Ness, Jeffrey
APPLICANT: Winkler, David G.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
FILE REFERENCE: 240083.508
CURRENT APPLICATION NUMBER: US/09/449,218D
CURRENT FILING DATE: 1999-11-24
NUMBER OF SEQ ID NOS: 45
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 27
LENGTH: 57
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE: Primer for PCR
OTHER INFORMATION: Primer for PCR
US-09-449-218D-27

```

Query Match	2.8%;	Score 21;	DB 4;	Length 57;
Best Local Similarity	100.0%;	Pred. No. 1.2;		
Matches 21;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY 655 GCCGAGCTGGAGAACGCTTAC 675
 |||||
 Db 57 GCCGAGCTGGAGAACGCTTAC 37

RESULT 15

US-09-213-767-1
 ; Sequence 1, Application US/09213767
 ; Patent No. 5948680
 ; GENERAL INFORMATION:
 ; APPLICANT: Brenda F. Baker
 ; APPLICANT: Lex M. Cowsett
 ; TITLE OF INVENTION: ANTISENSE MODULATION OF ELK-1 EXPRESSION
 ; FILE REFERENCE: RTS-0024
 ; CURRENT APPLICATION NUMBER: US/09/213,767
 ; CURRENT FILING DATE: 1998-12-17
 ; NUMBER OF SEQ ID NOS: 47
 ; SEQ ID NO 1
 ; LENGTH: 2266
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (316)..(1602)
 US-09-213-767-1

Query Match 2.8%; Score 21; DB 2; Length 2266;
 Best Local Similarity 100.0%; Pred. No. 0.98;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 600 GCCGCAAGAGGCGCGAAGCC 620
 |||||
 Db 1242 GCCGCAAGAGGCGCGAAGCC 1262

Search completed: March 29, 2003, 10:46:53
 JOD time : 66.7914 secs


```

QY 241 CCCTTTGAGACAAAGAGCTGTCCGAGTACAGTCCGCGAGCTGACTTACCCGCTAC 300
D 241 CCCTTTGAGACAAAGAGCTGTCCGAGTACAGTCCGCGAGCTGACTTACCCGCTAC 300
QY 301 GTGACGATGAGGCGGTGCGCGAGCGCAAGCGGTGCAACGAGTGTGTCTCCGCGCAG 360
D 301 GTGACGATGAGGCGGTGCGCGAGCGCAAGCGGTGCAACGAGTGTGTCTCCGCGCAG 360
QY 361 TGGCGGCGGCGGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 420
D 361 TGGCGGCGGCGGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 420
QY 421 GGGCGCGGACTTCCGCTGATCCCGAGCGCTACCGCGCGAGCGCGGAGCTGTGTGTGT 480
D 421 GGGCGCGGACTTCCGCTGATCCCGAGCGCTACCGCGCGAGCGCGGAGCTGTGTGTGT 480
QY 481 CCCGCTGTGAGAGCGCGCGCGCGCGAGTGGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 540
D 481 CCCGCTGTGAGAGCGCGCGCGCGCGAGTGGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 540
QY 541 CGCTCTACCGCTTCTCCACACAGTGGAGTCAAGAGCTTGGGAGCCGAGCGCTCGG 600
D 541 CGCTCTACCGCTTCTCCACACAGTGGAGTCAAGAGCTTGGGAGCCGAGCGCTCGG 600
QY 601 CCGGAGAGGCGCGAGAGCGCGCGCGCGCGCGAGCGCGCAAGCGCAACGAGCGGAG 660
D 601 CCGGAGAGGCGCGAGAGCGCGCGCGCGCGCGAGCGCGCAAGCGCAACGAGCGGAG 660
QY 661 CTGAGAGAGCGCTTCTGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 720
D 661 CTGAGAGAGCGCTTCTGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 720
QY 721 GAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 759
D 721 GAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 759

```

RESULT 2

```

US-09-864-761-30988/c
; Sequence 30988, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668

```

```

; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
; SEQ ID NO 30988
; LENGTH: 422
; TYPE: DNA
; ORGANISM: Homo sapiens
; OTHER INFORMATION: MAP TO AC003098.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.62
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.6
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.49
; OTHER INFORMATION: BEST HUMAN HIT: BE613498.1, EVALUATE 9.90e-02
; OTHER INFORMATION: SWISSPROT HIT: P45646, EVALUATE 4.90e-01
; OTHER INFORMATION: NT HIT: AF074705.1, EVALUATE 1.00e+00
US-09-864-761-30988

```

```

Query Match 55.6%; Score 422; DB 10; Length 422;
Best Local Similarity 100.0%; Pred. No. 6e-214;
Matches 422; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 257 ACGTGTCCGAGTACAGCTGCGCGAGCTGCACTTCAACCGCTACGATGAGCGCT 316
D 422 ACGTGTCCGAGTACAGCTGCGCGAGCTGCACTTCAACCGCTACGATGAGCGCT 363
QY 317 GCCGAGGCGCCAAAGCGGTGCAACGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 376
D 362 GCCGAGGCGCCAAAGCGGTGCAACGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 303
QY 377 TCGTGTCCCAAGCCATCGAGCGCGCGCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 436
D 302 TCGTGTCCCAAGCCATCGAGCGCGCGCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 243
QY 437 GCATCCCGGACCGCTACCGCGCGAGCGCGGTGCAAGCTGTGTGTGTGTGTGTGTGTGT 496
D 242 GCATCCCGGACCGCTACCGCGCGAGCGCGGTGCAAGCTGTGTGTGTGTGTGTGTGTGT 183
QY 497 CGCGGCGCGGCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 556
D 182 CGCGGCGCGGCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 123
QY 557 ACAACCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 616
D 122 ACAACCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 63
QY 617 AGCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTACT 676
D 62 AGCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTACT 3
QY 677 AG 678
D 2 AG 1

```

RESULT 3

```

US-09-864-761-14440/c
; Sequence 14440, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:

```

```

APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DEIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aemita-X-1
CURRENT FILING DATE: US/09/864,761
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 14440
LENGTH: 392
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC003098.1
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.62
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.6
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.49
US-09-864-761-14440

Query Match          46.8%; Score 355; DB 10; Length 392;
Best Local Similarity 100.0%; Pred. No. 2e-178;
Matches 355; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

585 GACCGAGCCCTCGGCGGAGAGAGGCGGAGCCGCGCCCGCGGAGCGCCAA 644
212 GACCGAGCCCTCGGCGGAGAGAGGCGGAGCCGCGCCCGCGGAGCGCCAA 153
645 AGCCACCAAGGCGGAGGCTGAGAGAGGCTTAAAGCCCGCGCGCCCTTCCACCGG 704
152 AGCCACCAAGGCGGAGGCTGAGAGAGGCTTAAAGCCCGCGCGCCCTTCCACCGG 93
705 GCGGCGCCCGCGGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 759
92 GCGGCGCCCGCGGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 38

RESULT 4
US-09-867-274-3
Sequence 3, Application US/09867274
Patent No. US20020106650A1
GENERAL INFORMATION:
APPLICANT: Paszty, Christopher
TITLE OF INVENTION: Cysteine Knot Polypeptides: Cloaked-2 Molecules and Uses Thereof
FILE REFERENCE: 01017/37428
CURRENT FILING DATE: US/09/867,274
PRIOR APPLICATION NUMBER: US 60/208,550
PRIOR FILING DATE: 2000-06-01
PRIOR APPLICATION NUMBER: US 60/223,542
PRIOR FILING DATE: 2000-08-04
NUMBER OF SEQ ID NOS: 25
SOFTWARE: PatentIn version 3.0
SEQ ID NO 3
LENGTH: 636
TYPE: DNA
ORGANISM: Mus musculus
US-09-867-274-3

Query Match          9.1%; Score 69; DB 10; Length 636;
Best Local Similarity 100.0%; Pred. No. 7.4e-27;
Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

517 CTGTGGCTCTGTGCAAGTGAAGGCGGCTTCCAGCAACGAGTGGAGCTTAAG 576
475 CTGTGGCTCTGTGCAAGTGAAGGCGGCTTCCAGCAACGAGTGGAGCTTAAG 534
577 GACTTCGGG 585
535 GACTTCGGG 543

RESULT 5
US-09-867-274-21
Sequence 21, Application US/09867274
Patent No. US20020106650A1
GENERAL INFORMATION:
APPLICANT: Paszty, Christopher
TITLE OF INVENTION: Cysteine Knot Polypeptides: Cloaked-2 Molecules and Uses Thereof
FILE REFERENCE: 01017/37428
CURRENT FILING DATE: US/09/867,274
PRIOR APPLICATION NUMBER: US 60/208,550
PRIOR FILING DATE: 2000-06-01
PRIOR APPLICATION NUMBER: US 60/223,542
PRIOR FILING DATE: 2000-08-04
NUMBER OF SEQ ID NOS: 25
SOFTWARE: PatentIn version 3.0
SEQ ID NO 21
LENGTH: 27
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Artificial. PCR primer

```



```
RESULT 10
US-09-815-242-5978/c
; Sequence 5978, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5978
; LENGTH: 645
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(645)
US-09-815-242-5978

Query Match          2.6%; Score 20; DB 10; Length 645;
Best Local Similarity 100.0%; Pred. No. 0.7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 625 CCCCCGGCCGAGGCCCA 644
DB 36 CCCCCGGCCGAGGCCCA 17

RESULT 11
US-09-815-242-7536/c
; Sequence 7536, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9619
; LENGTH: 645
; TYPE: DNA
; ORGANISM: Salmonella typhi
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(645)
```

```

; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7536
; LENGTH: 645
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(645)
US-09-815-242-7536

Query Match          2.5%; Score 19; DB 10; Length 645;
Best Local Similarity 100.0%; Pred. No. 2.4;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 626 CCCCCGGCCGAGGCCCA 644
DB 35 CCCCCGGCCGAGGCCCA 17

RESULT 12
US-09-815-242-9619/c
; Sequence 9619, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9619
; LENGTH: 645
; TYPE: DNA
; ORGANISM: Salmonella typhi
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(645)
```

US-09-815-242-9619

Query Match 2.5%; Score 19; DB 10; Length 645;
Best Local Similarity 100.0%; Pred. No. 2.4;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 626 CCGCGCCGAGAGCCCA 644
DB 35 CCGCGCCGAGAGCCCA 17

RESULT 13

US-09-903-376-1
; Sequence 1, Application US/09903376
; Publication No. US20030023998A1
; GENERAL INFORMATION:
; APPLICANT: Brennan, Thomas J.
; TITLE OF INVENTION: TRANSGENIC MICE CONTAINING 5-HT-2B GENE
; FILE REFERENCE: R-599
; CURRENT APPLICATION NUMBER: US/09/903,376
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/218,358
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/223,120
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 60/223,122
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1550
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-903-376-1

Query Match 2.5%; Score 19; DB 9; Length 1550;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 43 CTCCTGCTGCTGCTGCTC 61
DB 399 CTCCTGCTGCTGCTGCTC 417

RESULT 14

US-09-867-274-10/c
; Sequence 10, Application US/09867274
; Patent No. US20020106650A1
; GENERAL INFORMATION:
; APPLICANT: Paazty, Christopher
; TITLE OF INVENTION: Cysteine Knot Polypeptides: Cloaked-2 Molecules and Uses Thereof
; FILE REFERENCE: 01017/37428
; CURRENT APPLICATION NUMBER: US/09/867,274
; PRIOR FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: US 60/208,550
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: US 60/223,542
; PRIOR FILING DATE: 2000-08-04
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 10
; LENGTH: 24
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial: PCR primer
US-09-867-274-10

Query Match 2.4%; Score 18; DB 10; Length 24;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 426 CGACTTCGCTGATCCC 443
DB 24 CGACTTCGCTGATCCC 7

RESULT 15

US-09-864-761-3211/c
; Sequence 3211, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aesomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 3211
; LENGTH: 375
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL121752.2
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 22
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.4
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 4.5
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN HEL100, SIGNAL = 8.5
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 4.7
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.8

; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 7.2
US-09-864-761-3211

Query Match 2.4%; Score 18; DB 10; Length 375;
Best Local Similarity 100.0%; Pred. No. 8.5;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 464 GCCTGCAGCTGCTGTGTC 481
|||
Db 213 GCCTGCAGCTGCTGTGTC 196

Search completed: March 29, 2003, 10:49:31
Job time : 94.4064 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 29, 2003, 01:29:00 ; Search time 1634.43 Seconds
(without alignments) 7520.874 Million cell updates/sec

Title: US-09-867-274-1
Page: 750

Sequence: 1 tactggaagtcgtgcc.....ctgtcctctgcgcgtgtt 759

Scoring table: OLIGO_NUC

Searched: 16154066 beqs, 8097743376 residues

Word size :

Total number of hits satisfying chosen parameters: 32308132

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Listing first 45 summaries

Database

```

1: em.estbta:*
2: em.estbta:*
3: em.estbta:*
4: em.estbta:*
5: em.estbta:*
6: em.estbta:*
7: em.estbta:*
8: em.estbta:*
9: em.estbta:*
10: em.estbta:*
11: em.estbta:*
12: em.estbta:*
13: em.estbta:*
14: em.estbta:*
15: em.estbta:*
16: em.estbta:*
17: em.estbta:*
18: em.estbta:*
19: em.estbta:*
20: em.estbta:*
21: em.estbta:*
22: em.estbta:*
23: em.estbta:*
24: em.estbta:*
25: em.estbta:*
26: em.estbta:*
27: em.estbta:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	232	30.6	535	17	A0171546	A0171546 HS 3088 F
C 2	69	9.1	623	10	B8637315	B8637315 BB637315
C 3	69	9.1	1990	11	AK017295	AK017295 Mus musc
C 4	68	9.0	291	10	BE111224	BE111224 U1-R-BJ1
C 5	68	9.0	360	9	A1565282	A1565282 U1-R-C2P
C 6	68	9.0	419	10	BE101082	BE101082 U1-R-BJ1

[illegible]

REFERENCE	1 (baaes 1 to 535)
AUTHORS	Mahitres,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,U., Zhao,S., Adams,M.D. and Hood,L.
TITLE	Sequence-tagged connectors: A sequence approach to mapping and
JOURNAL	examining the human genome
MEDLINE	Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
COMMENT	99380589 Contact: Mahitres GG, Wallace JC, Hood L

Contact: Mahatras GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel.: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Sequence Tagged Connector
Plate: 3088 row: P column: 11

ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE 1 (bases 1 to 291)
 AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene
 discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 COMMENT Contact: Soares, MB
 Program for Rat Gene Discovery and Mapping
 University of Iowa
 451 Eckstein Medical Research Building Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: msoares@blue.weeg.uiowa.edu
 Oligo-dT track not found, Not 1 site shown in beginning of sequence
 is likely internal to the message. cDNA library Preparation: M.B.
 Soares Lab Clone distribution: clones will be available through
 Research Genetics (www.resgen.com) The following repetitive
 elements were found in this cDNA sequence: 97-125,
 >GC_rich#low complexity
 Seq primer: M13 Forward
 POLYA=No.

FEATURES
 source Location/Qualifiers
 1..291
 /organism="Rattus norvegicus"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /clone="UI-R-BJ1-auw-b-04-0-UI"
 /clone_1lb="UI-R-BJ1"
 /lab_host="DH10B (Life Technologies)"
 /note="Vector: pRT3D-Pac (Pharmacia) with a modified
 polylinker; Site 1: Not 1; Site 2: Eco RI; The UI-R-BJ1
 library is a subtracted library derived from the following
 tissues: atrium at 16.5 dpc, ventricle at 16.5 dpc, AV
 canal at 16.5 dpc, atrium at 15 dpc, ventricle at 15 dpc,
 AV canal at 15 dpc, ventricle at 13 dpc, and adult heart.
 For a detailed description of the library from which this
 clone was derived, please visit our web site at
 rsearch.eng.uiowa.edu. The subtraction has been previously
 described in (Bonaldo, Lennon and Soares, Genome Research
 6:791-806, 1996)
 TAG_SEQ=None found"
 BASE COUNT 41 a 101 c 109 g 40 t

ORIGIN
 Query Match 9.0%; Score 68; DB 10; Length 291;
 Best Local Similarity 100.0%; Pred. No. 3.3e-22;
 Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 517 CTGTGGCTCTGTGCAAGTGCAGCGCTTCCACCAACGAGCTCAAG 576
 |||||
 DB 87 CTGTGGCTCTGTGCAAGTGCAGCGCTTCCACCAACGAGCTCAAG 28

QY 577 GACTTCGG 584
 |||||
 DB 27 GACTTCGG 20

RESULT 5
 LOCUS AI556282/c 360 bp mRNA linear EST 23-MAR-1999
 DEFINITION UI-R-C2P-rk-a-03-0-UI.s1 UI-R-C2P Rattus norvegicus cDNA clone
 accession AI556282
 version AI556282.1 GI:4488645
 keywords EST.
 SOURCE Norway rat.
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

REFERENCE 1 (bases 1 to 360)
 AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene
 discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 COMMENT Contact: Soares, MB
 Program for Rat Gene Discovery and Mapping
 University of Iowa
 451 Eckstein Medical Research Building Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: msoares@blue.weeg.uiowa.edu
 Oligo-dT track not found, Not 1 site shown in beginning of sequence
 is likely internal to the message. cDNA Library Preparation: M.
 Fatima Bonaldo, Ph.D. Clone distribution: clones will be available
 through Research Genetics (www.resgen.com) The following repetitive
 elements were found in this cDNA sequence: 97-125,
 >GC_rich#low complexity
 Seq primer: M13 Forward

FEATURES
 source Location/Qualifiers
 1..360
 /organism="Rattus norvegicus"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /clone="UI-R-C2P-rk-a-03-0-UI"
 /clone_1lb="UI-R-C2P"
 /dev_stage="adult"
 /lab_host="DH10B (Life Technologies)"
 /note="Vector: pRT3D-Pac (Pharmacia) with a modified
 polylinker; Site 1: Not 1; Site 2: Eco RI; The UI-R-C2P
 library is a subtracted library derived from the UI-R-C1
 library, which is a subtracted library derived from the
 mixture of individually tagged normalized libraries
 constructed from rat placenta, adult lung, brain, liver,
 kidney, heart, spleen, ovary, muscle, 8, 12 and 18-day
 embryo. The tag is a string of 3-5 nucleotides present
 between the Not I site and the oligo-dT track which allows
 identification of the library of origin of a clone within
 the mixture. The subtracted library (UI-R-C2P) was
 constructed as follows: PCR amplified cDNA inserts from
 UI-R-C1 clones from which 3' ESTs had been derived was
 used as a driver in a hybridization with the UI-R-C1
 library in the form of single-stranded circles. The
 remaining single-stranded circles (subtracted library) was
 purified by hydroxyapatite column chromatography,
 converted to double-stranded circles and electroporated
 into DH10B bacteria (Life Technologies) to generate the
 UI-R-C2P library. This procedure has been previously
 described (Bonaldo, Lennon and Soares, Genome Research 6:
 791-806, 1996)"

BASE COUNT 54 a 120 c 131 g 54 t 1 others

ORIGIN
 Query Match 9.0%; Score 68; DB 9; Length 360;
 Best Local Similarity 100.0%; Pred. No. 3.5e-22;
 Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 517 CTGTGGCTCTGTGCAAGTGCAGCGCTTCCACCAACGAGCTCAAG 576
 |||||
 DB 87 CTGTGGCTCTGTGCAAGTGCAGCGCTTCCACCAACGAGCTCAAG 28

QY 577 GACTTCGG 584
 |||||
 DB 27 GACTTCGG 20

RESULT 6
 LOCUS BE101082/c 419 bp mRNA linear EST 13-JUN-2000
 DEFINITION UI-R-BJ1-ety-h-10-0-UI.s1 UI-R-BJ1 Rattus norvegicus cDNA clone

ACCESSION UI-R-BJ1-aty-h-10-0-UI 3', mRNA sequence.
 VERSION BE101082
 KEYWORDS BE101082.1 GI:8492989
 SOURCE EST
 ORGANISM Norway rat.
 Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 419)
 REFERENCE Bonaldo,M.F., Lennon,G. and Soares,M.B.
 AUTHORS Normalization and subtraction: two approaches to facilitate gene
 TITLE discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 COMMENT Contact: Soares, MB
 Program for Rat Gene Discovery and Mapping
 University of Iowa
 451 Eckstein Medical Research Building Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: msoares@blue.weeg.uiowa.edu
 Oligo-dT track not found, Not 1 site shown in beginning of sequence
 is likely internal to the message. cDNA Library Preparation: M.B.
 Soares Lab Clone distribution: clones will be available through
 Research Genetics (www.resgen.com) The following repetitive
 elements were found in this cDNA sequence: 97-125,
 >GC_rich#low_complexity
 Seq primer: M13 Forward
 POLYA=No.

FEATURES
 source
 1..419 Location/Qualifiers
 /organism="Rattus norvegicus"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /clone="UI-R-BJ1-aty-h-10-0-UI"
 /clone_1ib="UI-R-BJ1"
 /lab_host="DH10B (Life Technologies)"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker. Site_1: Not 1; Site_2: Eco RI; The UI-R-BJ1
 library is a subtracted library derived from the following
 tissues: atrium at 16.5 dpc, ventricle at 16.5 dpc, AV
 canal at 16.5 dpc, atrium at 15 dpc, ventricle at 15 dpc,
 AV canal at 15 dpc, ventricle at 13 dpc, and adult heart.
 For a detailed description of the library from which this
 clone was derived, please visit our web site at
 ratest.eng.uiowa.edu. The subtraction has been previously
 described in (Bonaldo, Lennon and Soares, Genome Research
 6:791-806, 1996)
 TAG_SEQ=None found"
 BASE COUNT 56 a 135 c 155 g 73 t
 ORIGIN

Query Match 9.0%; Score 68; DB 10; Length 419;
 Best Local Similarity 100.0%; Pred. No. 3.5e-22;
 Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 517 CTGTGGCTCTGTGCAAGTGCAGCGCTTCACACCAAGTGCAGCTCAAG 576
 |||||||
 DB 87 CTGTGGCTCTGTGCAAGTGCAGCGCTTCACACCAAGTGCAGCTCAAG 28
 |||||||
 QY 577 GACTTCGG 584
 |||||||
 DB 27 GACTTCGG 20
 |||||||

RESULT 7
 A1113131/c 315 bp mRNA linear EST 11-FEB-1999
 LOCUS UI-R-C2p-nr-f-06-0-UI.81 UI-R-C2p Rattus norvegicus cDNA clone
 DEFINITION UI-R-C2p-nr-f-06-0-UI 3', mRNA sequence.
 A1113131
 ACCESSION A1113131.1 GI:3513080
 VERSION

KEYWORDS EST.
 SOURCE Norway rat.
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 315)
 REFERENCE Bonaldo,M.F., Lennon,G. and Soares,M.B.
 AUTHORS Normalization and subtraction: two approaches to facilitate gene
 TITLE discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 COMMENT Contact: Soares, MB
 Program for Rat Gene Discovery and Mapping
 University of Iowa
 451 Eckstein Medical Research Building Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: msoares@blue.weeg.uiowa.edu
 Oligo-dT track not found, Not 1 site shown in beginning of sequence
 is likely internal to the message. cDNA Library Preparation: M.
 Fatima Bonaldo, Ph.D. Clone distribution: clones will be available
 through Research Genetics The following repetitive elements were
 found in this cDNA sequence: 97-125, >GC_rich#low_complexity
 Seq primer: M13 Forward.

FEATURES
 source
 1..315 Location/Qualifiers
 /organism="Rattus norvegicus"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /clone="UI-R-C2p-nr-f-06-0-UI"
 /clone_1ib="UI-R-C2p"
 /dev_stage="adult"
 /lab_host="DH10B (Life Technologies)"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker. Site_1: Not 1; Site_2: Eco RI; The UI-R-C2p
 library is a subtracted library derived from the UI-R-C1
 library, which is a subtracted library derived from the
 UI-R-C0 library. The UI-R-C0 library consisted of a
 mixture of individually tagged normalized libraries
 constructed from rat placenta, adult lung, brain, liver,
 kidney, heart, spleen, ovary, muscle, 8, 12 and 18-day
 embryo. The tag is a string of 3-5 nucleotides present
 between the Not I site and the oligo-dT track which allows
 identification of the library of origin of a clone within
 the mixture. The subtracted library (UI-R-C2p) was
 constructed as follows: PCR amplified cDNA inserts from
 UI-R-C1 clones from which 3' ESTs had been derived was
 used as a driver in a hybridization with the UI-R-C1
 library in the form of single-stranded circles. The
 remaining single-stranded circles (subtracted library) was
 purified by hydroxyapatite column chromatography,
 converted to double-stranded circles and electroporated
 into DH10B bacteria (Life Technologies) to generate the
 UI-R-C2p library. This procedure has been previously
 described (Bonaldo, Lennon and Soares, Genome Research 6:
 791-806, 1996)"

BASE COUNT 46 a 110 c 116 g 43 t
 ORIGIN

Query Match 5.0%; Score 38; DB 9; Length 315;
 Best Local Similarity 100.0%; Pred. No. 1.3e-07;
 Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 518 TGTGTGGCTCTGTGCAAGTGCAGCGCTTCACCGCTTC 555
 |||||||
 DB 86 TGTGTGGCTCTGTGCAAGTGCAGCGCTTCACCGCTTC 49
 |||||||

RESULT 8
 BB638050 511 bp mRNA linear EST 31-AUG-2001
 LOCUS BB638050
 DEFINITION BB638050 RIKEN full-length enriched, 3 days neonate thymus Mus

Seq primer: SP6
Class: BAC ends.

FEATURES

Location/Qualifiers

1..147
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-24-79P16"
/sex="Male"
/cell_type="Spleen/Brain"
/note="Vector: pTRABACT; Site 1: BamHI, Site 2: BamHI, RPCI-24 Mouse BAC library produced by Pieter de Jong. The library was cloned in the pTRABACT cloning vector at the BamHI sites using MboI partially digested male C57BL/6J DNA."

BASE COUNT 30 a 44 c 45 g 28 t

ORIGIN

Query Match 3.4%; Score 26; DB 17; Length 147;

Best Local Similarity 100.0%; Pred. No. 0.077; Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 264 CGAGTACAGCTGCCGCGAGCTGCACT 289
DB 101 CGAGTACAGCTGCCGCGAGCTGCACT 126

RESULT 12
BG711211 368 bp mRNA linear EST 08-MAY-2001
LOCUS p911n.pk006.114 Normalized Liver library Gallus gallus cDNA clone
DEFINITION p911n.pk006.114 5' similar to g1485130 ref|NM_005195.1| Homo sapiens CCAT/enhancer binding protein (C/EBP), delta (CEBPd), mRNA, mRNA sequence.

ACCESSION BG711211 GI:14005161
VERSION BG711211.1
KEYWORDS EST.
SOURCE chicken.
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianidae; Gallus.
1 (bases 1 to 368)
Burnside, J., Morgan, R.W. and Cogburn, L.A.
Chicken ESTs from a normalized liver library
Unpublished (2001)
Contact: Joan Burnside
Molecular Endocrinology
University of Delaware
40 Townsend Hall, Newark, DE 19717, USA
Tel: 302 831-1345
Fax: 302 831-3411
Email: joan@udel.edu, www.chickest.udel.edu.

FEATURES
source
1..368
Location/Qualifiers

/organism="Gallus gallus"
/db_xref="taxon:9031"
/clone="p911n.pk006.114"
/clone_lib="Normalized Liver Library"
/sex="Male and Female"
/tissue_type="liver"
/lab_host="E.coli EMDH10B"
/note="Vector: PCMVSPORT 6"

BASE COUNT 74 a 127 c 102 g 65 t

ORIGIN

Query Match 3.3%; Score 25; DB 12; Length 368;
Best Local Similarity 100.0%; Pred. No. 0.28;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 618 GCCGCGGCCGCCGCCGCGAGCGCC 642

DB 233 GCCGCGGCCGCCGCCGCGAGCGCC 257

RESULT 13

LOCUS AW434705 385 bp mRNA linear EST 11-FEB-2000
DEFINITION UI-R-BJ0P-afz-c-02-0-UI.61 UI-R-BJ0P Rattus norvegicus cDNA clone
UI-R-BJ0P-afz-c-02-0-UI 3', mRNA sequence.
ACCESSION AW434705
VERSION AW434705.1 GI:6966012
KEYWORDS EST.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 385)
Bonaldo, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene discovery
Genome Res. 6 (9), 791-806 (1996)
97044477
Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: msoares@iue.uiowa.edu

COMMENT
Oligo-dt track not found. Not 1 site shown in beginning of sequence
is likely internal to the message. cDNA library preparation: M.B.
Soares Lab Clone distribution: clones will be available through
Research Genetics (www.resgen.com) The following repetitive
elements were found in this cDNA sequence: 21-41,
xgc_richlow_complexity
Seq primer: M13 forward
POLYA=NO.

FEATURES

source

1..385
Location/Qualifiers
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-BJ0P-afz-c-02-0-UI"
/clone_lib="UI-R-BJ0P"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies)"
/note="Vector: p773D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-BJ0P library is a subtracted library derived from the UI-R-A1, UI-R-A1, UI-R-A1, UI-R-A1, UI-R-A1, UI-R-A1, and UI-R-A1 libraries. These libraries represent tissues from rat atrium at 16.5 dpc, ventricle at 16.5 dpc, AV canal at 16.5 dpc, atrium at 15 dpc, ventricle at 15 dpc, AV canal at 15 dpc, and ventricle at 13 dpc. The tag is a string of 5-6 nucleotides present between the Not I site and the oligo-dt track. The library was constructed as described by Bonaldo, Lennon and Soares, Genome Research 6: 791-806, 1996.
TAG_LIB=UI-R-BJ0P
TAG_TISSUE=AV canal at 16.5 dpc
TAG_SEQ=GAACCC

BASE COUNT 75 a 127 c 122 g 61 t

ORIGIN

Query Match 2.8%; Score 21; DB 10; Length 385;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 616 AAGCCGCGGCCGCCGCCGCGG 636
DB 19 AAGCCGCGGCCGCCGCCGCGG 39

```

RESULT 14
BF420024          472 bp  mRNA  linear  EST 28-NOV-2000
LOCUS             UI-R-BJ2-bpm-b-02-0-UI. s1 UI-R-BJ2 Rattus norvegicus cDNA clone
DEFINITION       UI-R-BJ2-bpm-b-02-0-UI 3', mRNA sequence.
ACCESSION       BF420024
VERSION         BF420024.1  GI:11408013
KEYWORDS
SOURCE
ORGANISM        Rattus norvegicus
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
                Rattus.
REFERENCE       1 (bases 1 to 472)
AUTHORS        Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE          Normalization and subtraction: two approaches to facilitate gene
                discovery
JOURNAL        Genome Res. 6 (9), 791-806 (1996)
MEDLINE        97044477
COMMENT        Contact: Soares, MB
                Program for Rat Gene Discovery and Mapping
                University of Iowa
                451 Eckstein Medical Research Building Iowa City, IA 52242, USA
                Tel: 319 335 8250
                Fax: 319 335 9565
                Email: msoares@blue.weeg.uiowa.edu
                Oligo-dt track not found. Not 1 site shown in beginning of sequence
                is likely internal to the message. cDNA library Preparation: M.B.
                Soares Lab Clone distribution: clones will be available through
                Research Genetics (www.resgen.com) The following repetitive
                elements were found in this cDNA sequence: 22-42,
                >GC rich#Low complexity
                Seq primer: M13 Forward
                POLYA=No.
FEATURES
source
1..472
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone_lib="UI-R-BJ2-bpm-b-02-0-UI"
/lab_host="DH10B (Life Technologies)"
/notes="Vector: pTRF3D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-BJ2
library is a subtracted library derived from the following
tissues: heart, atrium at 15 dpc, ventricle at 16.5 dpc,
atrium at 16.5 dpc, ventricle at 13 dpc, ventricle at 15
dpc, AV canal at 15 dpc. For a detailed description of
the library from which this clone was derived, please
visit our web site at ratest.eng.uiowa.edu. The
subtraction has been previously described in (Bonaldo,
Lennon and Soares, Genome Research 6:791-806, 1996)
TMG SRO=None found"
BASE COUNT      103 a 148 c 147 g 74 t
ORIGIN
Query Match      2.8%; Score 21; DB 12; Length 472;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 616 AAGCCGCGCGCCCGCGCCGCG 636
DB 20 AAGCCGCGCGCCCGCGCCGCG 40
RESULT 15
AQ803172          477 bp  DNA  linear  GSS 09-AUG-1999
LOCUS            HS 3158 B1 G02 T7C CIT Approved Human Genomic Sperm Library D Homo
DEFINITION       sapiens genomic clone Plate=3158 Col=3 Row=N, DNA sequence.
ACCESSION       AQ803172
VERSION         AQ803172.1  GI:5720504
KEYWORDS        GSS.

```

```

SOURCE          human.
ORGANISM        Homo sapiens
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE       1 (bases 1 to 477)
AUTHORS        Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
                Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
                Hood,L.
TITLE          Sequence-tagged connectors: A sequence approach to mapping and
                scanning the human genome
JOURNAL        Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
MEDLINE        99380589
COMMENT        Contact: Mahairas GG, Wallace JC, Hood L
                High Throughput Sequencing Center
                University of Washington
                401 Queen Anne Avenue North, Seattle, WA 98109, USA
                Tel: (206) 616-3618
                Fax: (206) 616-3887
                Email: jwallace@u.washington.edu
                Clones may be purchased from Research Genetics (info@resgen.com).
                BAC end Web Server: http://www.htsc.washington.edu
                Plate: 3158 row: N column: 3
                Seq primer: T7
                Class: BAC ends
                High quality sequence stop: 477.
FEATURES
source
1..477
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_plate="3158 Col=3 Row=N"
/clone_lib="CIT Approved Human Genomic Sperm Library D"
/sex="male"
/notes="Organ: sperm; Vector: pBelobAC11; BAC clones in
E-Coli DH10B"
BASE COUNT      76 a 171 c 143 g 82 t 5 others
ORIGIN
Query Match      2.8%; Score 21; DB 17; Length 477;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 600 GCCGCAAGAGGCGCGGAAGCC 620
DB 171 GCCGCAAGAGGCGCGGAAGCC 191
Search completed: March 29, 2003, 10:45:04
Job time : 1639.43 secs

```


GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 28, 2003, 14:06:06 ; Search time 53.2 Seconds
(without alignments)
475.895 Million cell updates/sec

Title: US-09-867-274-2
Parfact score: 1049

Sequence: 1 QGWOAFKNDATETIPELGEY.....KPPRARSAKANQAELENAY 190

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: /SIDS2/gcgdata/genseq/genseqp-emb1/AA1980.DAT.*

2: /SIDS2/gcgdata/genseq/genseqp-emb1/AA1981.DAT.*

3: /SIDS2/gcgdata/genseq/genseqp-emb1/AA1982.DAT.*

4: /SIDS2/gcgdata/genseq/genseqp-emb1/AA1983.DAT.*

5: /SIDS2/gcgdata/genseq/genseqp-emb1/AA1984.DAT.*

6: /SIDS2/gcgdata/genseq/genseqp-emb1/AA1985.DAT.*

7: /SIDS2/gcgdata/genseq/genseqp-emb1/AA1986.DAT.*

8: /SIDS2/gcgdata/genseq/genseqp-emb1/AA1987.DAT.*

9: /SIDS2/gcgdata/genseq/genseqp-emb1/AA1988.DAT.*

10: /SIDS2/gcgdata/genseq/genseqp-emb1/AA1989.DAT.*

11: /SIDS2/gcgdata/genseq/genseqp-emb1/AA1990.DAT.*

12: /SIDS2/gcgdata/genseq/genseqp-emb1/AA1991.DAT.*

13: /SIDS2/gcgdata/genseq/genseqp-emb1/AA1992.DAT.*

14: /SIDS2/gcgdata/genseq/genseqp-emb1/AA1993.DAT.*

15: /SIDS2/gcgdata/genseq/genseqp-emb1/AA1994.DAT.*

16: /SIDS2/gcgdata/genseq/genseqp-emb1/AA1995.DAT.*

17: /SIDS2/gcgdata/genseq/genseqp-emb1/AA1996.DAT.*

18: /SIDS2/gcgdata/genseq/genseqp-emb1/AA1997.DAT.*

19: /SIDS2/gcgdata/genseq/genseqp-emb1/AA1998.DAT.*

20: /SIDS2/gcgdata/genseq/genseqp-emb1/AA1999.DAT.*

21: /SIDS2/gcgdata/genseq/genseqp-emb1/AA2000.DAT.*

22: /SIDS2/gcgdata/genseq/genseqp-emb1/AA2001.DAT.*

23: /SIDS2/gcgdata/genseq/genseqp-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1049	100.0	190	23	ABB07207	Human cloaked-2 po
2	1049	100.0	213	21	AAB26106	Human DAN/Cerberus
3	1049	100.0	213	21	AAV96429	Human TGF-beta bin
4	1049	100.0	213	22	AAV97589	Human secreted pro
5	1049	100.0	213	23	ABG34061	Human Pro peptide
6	1049	100.0	213	23	ABB07209	Human cloaked-2 po
7	1049	100.0	213	23	AAE17089	Human osteolevin p
8	1040	99.1	213	21	AAV96430	Human TGF-beta bin
9	1040	99.1	213	21	AAV96436	Human TGF-beta bin
10	1033	98.5	213	21	AAV96431	Verret TGF-beta bi

11	974	92.9	317	21	AA966431	Rat TGF-beta bindi
12	954	90.9	263	21	AA826105	Human DAN/Cerberus
13	951	90.7	188	23	AB807208	Mouse cloaked-2 po
14	951	90.7	211	23	AA936432	Murine TGF-beta b1
15	951	90.7	211	23	AB807210	Mouse cloaked-2 po
16	901.5	85.9	176	21	AA956434	Bovine TGF-beta b1
17	758	72.3	139	22	AA973067	Human bone marrow
18	758	72.3	139	23	ABG42911	Human peptide encocc
19	344	32.8	206	21	AA975981	Murine skin cell p
20	344	32.8	206	21	AA976031	Murine skin cell p
21	344	32.8	206	22	AA855920	Skin cell protein,
22	344	32.8	206	22	AA855970	Skin cell protein,
23	344	32.8	206	23	AB872120	Murine protein iso
24	344	32.8	206	23	AB872120	Murine protein iso
25	341	32.5	206	18	AA905948	Human small CCN-11
26	341	32.5	206	19	AA958704	Human small CCN-11
27	341	32.5	206	20	AA955711	Homo sapiens fetal
28	341	32.5	206	21	AA810233	Human adult retina
29	341	32.5	206	23	AB855457	Human angiogenesis
30	341	32.5	206	23	AB809200	Human small CCN-11
31	341	32.5	206	23	AB809200	SCGP CNV family pr
32	341	32.5	206	23	AB884851	Human PROS12 prote
33	337	32.1	183	23	AB807213	Human cloaked-1 p
34	336.5	32.1	186	23	AA982653	Human breast tumo
35	324.5	30.9	102	23	AA982655	Human breast tumo
36	148	14.1	50	20	AA912009	Human 5' EST seque
37	114	10.9	116	18	AA947654	Secreted protein A
38	114	10.9	116	18	AA944050	Human secreted pro
39	104.5	10.0	102	20	AA930225	Amino acid sequenc
40	104.5	10.0	102	21	AA910277	Human fetal kidney
41	95.5	9.1	418	22	AA893353	Human protein segm
42	92	8.8	832	22	ABG31165	Novel human diagn
43	90.5	8.6	168	21	AA948014	Amino acid sequenc
44	89.5	8.5	184	19	AA969293	Human b57 protein
45	89.5	8.5	184	20	AA942173	Mouse DRM protein

ALIGNMENTS

RESULT 1
ABB07207
ID ABB07207 standard; Protein; 190 AA

AC ABB07207

DT 26-MAR-2002 (first entry)

Human cloaked-2 polypeptide mature protein sequence

KM Cloaked-2; cysteine knot motif; nephrotoxic; cardiatic; immunomodulator;
KM hepatocytic; antinflammatory; antithyroid; cytosatic; neuroprotective
KM antiischemic; hypotensive; antiarrhythmic; antiarteriosclerotic; muscular;
KM antiabietic; anorectic; gene therapy; cell therapy; antisense therapy;
KM human.

OS Homo sapiens

FH	Key	Location/Qualifiers
cm	3456789012	1c

FT /note="this region has been repeated twice in the
FT sequence provided in the sequence listing but has
FT been indicated correctly in the sequence in the
FT Figure"

PN WO200192308-A2.

PD 06-DEC-2001

PF 29-MAY-2001; 2001WO-US17478
YY

PR	01-JUN-2000; 2000US-208550P
PR	04-JUN-2000; 2000US-223543P

XX (AMGE-) AMGEN INC.
 PA Paszty CJ, Gao Y;
 PI
 XX WPI; 2002-114325/15.
 DR N-PSDB; ABA94293.
 XX
 PT New human and mouse cysteine-knot polypeptide designated as Cloaked-2,
 PT for treating or preventing kidney, heart (e.g. myocardial infarction)
 PT or liver (e.g. hepatitis) diseases
 XX
 PS Claim 13; Fig 1; 170pp; English.
 XX
 CC The invention relates to polypeptides comprising a cysteine knot motif
 CC and designated as Cloaked-2, derived from human and mouse. The cloaked-2
 CC polypeptides can be expressed by standard recombinant methodology. The
 CC cloaked-2 polynucleotides are useful in gene therapy and antisense
 CC therapy. The cloaked-2 polypeptides and polynucleotides are useful for
 CC treating, preventing, ameliorating or detecting diseases and disorders of
 CC the kidney (e.g. anemia, hypertension or low blood pressure), heart (e.g.
 CC cardiac hypertrophy, congestive heart failure, myocardial infarction,
 CC arrhythmias, atherosclerosis, hypertension or low blood pressure),
 CC skeletal muscle (e.g. muscular dystrophy or cachexia), placenta (e.g.
 CC congenital abnormalities or miscarriage), liver (e.g. hepatitis or
 CC cirrhosis), pancreas (e.g. diabetes or pancreatitis), thyroid (e.g.
 CC Grave's disease or myxedema) or adrenal cortex (e.g. Cushing's disease
 CC or Addison's disease), homeostasis or metabolic diseases (e.g. obesity,
 CC cancer or myopathies), infections, or autoimmune diseases. Selective
 CC binding agents may be used to modulate the biological activities of
 CC cloaked-2 polypeptides or to detect cloaked-2 polypeptide levels in a
 CC sample. Transgenic non-human animals are useful for drug candidate
 CC screening. The present sequence represents the human cloaked-2
 CC mature polypeptide sequence.
 CC
 XX
 SQ Sequence 190 AA;
 XX
 Query Match 100.0%; Score 1049; DB 23; Length 190;
 Best Local Similarity 100.0%; Pred. No. 1.5e-93;
 Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QGMOAFKNDATETIPELGYPEPPELENKTNRAENGSRPHHPETKDVSEYSCREL 60
 DB 1 QGMOAFKNDATETIPELGYPEPPELENKTNRAENGSRPHHPETKDVSEYSCREL 60
 QY 61 HFTRYVTDPGCRSAKPVTELVCSGCGPARLLPNAIGRKWMPBSGDPFCIPDPRYRQR 120
 DB 61 HFTRYVTDPGCRSAKPVTELVCSGCGPARLLPNAIGRKWMPBSGDPFCIPDPRYRQR 120
 QY 121 VOLCPGGEAPRARKVRLVASCCKRKLTRFNQSELDKFGTEAARPKGRKPRPARSAK 180
 DB 121 VOLCPGGEAPRARKVRLVASCCKRKLTRFNQSELDKFGTEAARPKGRKPRPARSAK 180
 QY 181 ANQAELENAY 190
 DB 181 ANQAELENAY 190
 OS Homo sapiens.

XX WO20055193-A2.
 PN 21-SEP-2000.
 PD
 XX 02-MAR-2000; 2000WO-US05537.
 PF
 XX 12-MAR-1999; 99US-0124118.
 PR
 XX (REGG-) REGENERON PHARM INC.
 PA
 XX Economides AN;
 PI
 XX WPI; 2000-638179/61.
 DR N-PSDB; AAA94051.
 XX
 PT Novel isolated, human DNA/Cerberus related protein 6 which include
 PT natural homologue, and polypeptides comprising DCR6 domain and nucleic
 PT acids encoding the proteins which are useful as probes and primers
 XX
 PS Claim 8; Fig 3; 40pp; English.
 XX
 CC The present sequence comprises the amino acid sequence encoded by exons 1
 CC and 4 of the human DAN/Cerberus-related protein 6 (hDCR6) coding
 CC sequence. The coding sequence was isolated from a human kidney cDNA
 CC library containing exons 1 and 4 of the sequence. hDCR6 is closely
 CC related to the DAN and DCR5 proteins, both of which act as antagonists of
 CC morphogenic proteins such as BMP. It is possible that the hDCR6 gene and
 CC protein can be used as immunogens, modulators of cell function, growth
 CC and differentiation, to reduce undesirable bone formation, to identify
 CC DCR6 binding agents, in diagnosis, and in gene therapy.
 CC
 XX
 SQ Sequence 213 AA;
 XX
 Query Match 100.0%; Score 1049; DB 21; Length 213;
 Best Local Similarity 100.0%; Pred. No. 1.7e-93;
 Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QGMOAFKNDATETIPELGYPEPPELENKTNRAENGSRPHHPETKDVSEYSCREL 60
 DB 24 QGMOAFKNDATETIPELGYPEPPELENKTNRAENGSRPHHPETKDVSEYSCREL 83
 QY 61 HFTRYVTDPGCRSAKPVTELVCSGCGPARLLPNAIGRKWMPBSGDPFCIPDPRYRQR 120
 DB 84 HFTRYVTDPGCRSAKPVTELVCSGCGPARLLPNAIGRKWMPBSGDPFCIPDPRYRQR 143
 QY 121 VOLCPGGEAPRARKVRLVASCCKRKLTRFNQSELDKFGTEAARPKGRKPRPARSAK 180
 DB 144 VOLCPGGEAPRARKVRLVASCCKRKLTRFNQSELDKFGTEAARPKGRKPRPARSAK 203
 QY 181 ANQAELENAY 190
 DB 204 ANQAELENAY 213
 OS Homo sapiens.
 OS
 OS 08-JUN-2000.

XX 24-NOV-1999; 99MO-US27990.
 PF 27-NOV-1998; 98US-0110283.
 XX
 XX (DARW-) DARWIN DISCOVERY LTD.
 PA
 PI Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepfer BW;
 PI Van Ness J, Winkler DG;
 DR N-PSDB; AAA29055.
 XX
 XX Nucleic acids (1) encoding a transforming growth factor beta binding
 PT protein, useful for identifying agents for treating osteopenia,
 PT osteoporosis and fractures
 XX
 PS Claim 2; Page 116; 162pp; English.
 XX
 XX This shows the human transforming growth factor-beta (TGF-beta)
 CC binding protein designated hBEER. The cDNA and protein may be used for
 CC prevention, treatment and diagnosis of diseases associated with
 CC inappropriate BEER expression. For example, they may be used to treat
 CC disorders associated with decreased TGF-beta BP expression. The cDNA or
 CC vectors may be administered to treat diseases by rectifying mutations or
 CC deletions in a patient's genome that affect the activity of BEER by
 CC expressing inactive proteins or to supplement the patient's own production
 CC of BEER polypeptides. The nucleic acids may be used for recombinant
 CC production of BEER, gene therapy, antisense therapy, as probes for
 CC diagnostic assays and for functional studies. BEER may be used to raise
 CC antibodies and for identification of BEER modulators. BEER antagonists
 CC may be used to increase bone mineral content for the treatment of
 CC disorders such as osteopenia, osteoporosis, fractures and other
 CC disorders associated with low mineral content.
 XX
 SQ Sequence 213 AA;
 Query Match 100.0%; Score 1049; DB 21; Length 213;
 Best Local Similarity 100.0%; Pred. No. 1.7e-93;
 Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 OGMQAFKNDATETIIPELGEYEPPEPELENNTKTNRAENGRRPHHPFTKDVSEYSCREL 60
 DB 24 OGMQAFKNDATETIIPELGEYEPPEPELENNTKTNRAENGRRPHHPFTKDVSEYSCREL 83
 QY 61 HFTRYVTDGCRSAKPYTELVCSCGCGPARLLPNAIGRKWMPBSGDPFRCTPDRYRAQR 120
 DB 84 HFTRYVTDGCRSAKPYTELVCSCGCGPARLLPNAIGRKWMPBSGDPFRCTPDRYRAQR 143
 QY 121 VOLLCPGGEAPARARVRLVASCKCKRLTRFNOSSELDKDFTEARPOKGRPRPARASAK 180
 DB 144 VOLLCPGGEAPARARVRLVASCKCKRLTRFNOSSELDKDFTEARPOKGRPRPARASAK 203
 QY 181 ANQAELENAY 190
 DB 204 ANQAELENAY 213
 RESULT 4
 AA97589
 ID AA97589 standard; Protein; 213 AA.
 AC AA97589;
 XX
 XX 05-APR-2001 (first entry)
 DT
 XX
 XX Human secreted protein PRO7476.
 DE
 XX
 XX Secreted protein; human; PRO protein; neoplastic cell growth; tumour;
 KW proliferation; leukaemia; lymphoid malignancy; inflammatory disorder;
 KW angiotonic disorder; immunologic disorder; PRO7476.
 XX
 OS Homo sapiens.

XX WO200075317-A2.
 PN
 XX
 PD 14-DEC-2000.
 XX
 PF 15-MAY-2000; 2000MO-US13358.
 XX
 XX 09-JUN-1999; 99US-0138385.
 PR 20-JUL-1999; 99US-0144790.
 PR 03-AUG-1999; 99US-0146843.
 PR 10-AUG-1999; 99US-0148188.
 PR 17-AUG-1999; 99US-0149320.
 PR 17-AUG-1999; 99US-0149327.
 PR 17-AUG-1999; 99US-0149396.
 PR 20-AUG-1999; 99US-0150114.
 PR 31-AUG-1999; 99US-0151700.
 PR 31-AUG-1999; 99US-0151734.
 XX
 PA (GETH) GENENTECH INC.
 XX
 PI Botstein DA, Goddard A, Gurney AL, Smith V, Watanabe CK, Wood WI;
 XX N-PSDB; AAA91023.
 DR
 DR WPI: 2001-071075/08.
 XX
 XX Antibodies against PRO polypeptides, useful for diagnosing and treating
 PT tumours are associated with gene amplification, neoplastic cell growth
 PT and proliferation in mammals -
 XX
 PS Claim 61; Fig 20; 143pp; English.
 XX
 XX This sequence is a human PRO protein of the invention. The PRO
 CC proteins are secreted proteins. Antagonists or antibodies of PRO
 CC polypeptides are useful for diagnosing and treating tumours are
 CC associated with gene amplification, neoplastic cell growth and
 CC proliferation in mammals, and those conditions characterised by
 CC overexpression and/or activation of the amplified genes. Such conditions
 CC include benign or malignant tumours (e.g. renal, liver, kidney, bladder,
 CC breast, gastric, ovarian, colorectal, prostate, pancreatic, lung, vulval,
 CC thyroid, hepatic carcinomas, sarcomas, glioblastomas and various head and
 CC neck tumours), leukaemias and lymphoid malignancies; neuronal, glial,
 CC astrocytal, hypochromic, and other glandular, macrophageal, epithelial,
 CC stromal and blastocoele disorders; and inflammatory, angiogenic and
 CC immunologic disorders. These may further be used to qualitatively or
 CC quantitatively detect the expression of proteins encoded by the
 CC amplified genes, and in tumour diagnostics or prognostics. The PRO
 CC polypeptide or its antagonist may be used for the preparation of a
 CC medicament in the treatment of a condition, which is responsive to the
 CC PRO polypeptide, its antagonist or anti-PRO antibody.
 XX
 SQ Sequence 213 AA;
 Query Match 100.0%; Score 1049; DB 22; Length 213;
 Best Local Similarity 100.0%; Pred. No. 1.7e-93;
 Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 OGMQAFKNDATETIIPELGEYEPPEPELENNTKTNRAENGRRPHHPFTKDVSEYSCREL 60
 DB 24 OGMQAFKNDATETIIPELGEYEPPEPELENNTKTNRAENGRRPHHPFTKDVSEYSCREL 83
 QY 61 HFTRYVTDGCRSAKPYTELVCSCGCGPARLLPNAIGRKWMPBSGDPFRCTPDRYRAQR 120
 DB 84 HFTRYVTDGCRSAKPYTELVCSCGCGPARLLPNAIGRKWMPBSGDPFRCTPDRYRAQR 143
 QY 121 VOLLCPGGEAPARARVRLVASCKCKRLTRFNOSSELDKDFTEARPOKGRPRPARASAK 180
 DB 144 VOLLCPGGEAPARARVRLVASCKCKRLTRFNOSSELDKDFTEARPOKGRPRPARASAK 203
 QY 181 ANQAELENAY 190
 DB 204 ANQAELENAY 213

RESULT 5
 ABG34061
 ID ABG34061 standard; Protein; 213 AA.
 AC ABG34061;
 XX
 DT 15-JUL-2002 (first entry)
 XX
 DE Human Pro peptide #32.
 XX
 KW Human; PRO; secreted protein; transmembrane protein;
 KW genetic disorder; tumour; cancer.
 XX
 OS Homo sapiens.
 XX
 PN WO200224888-A2.
 XX
 PD 28-MAR-2002.
 XX
 PF 29-AUG-2001; 2001WO-US27099.
 XX
 PR 01-SEP-2000; 2000US-229896P.
 PR 05-SEP-2000; 2000US-230621P.
 PR 22-SEP-2000; 2000US-235147P.
 PR 10-NOV-2000; 2000WO-US30873.
 PR 12-JAN-2001; 2001US-261878P.
 PR 16-JAN-2001; 2001US-261910P.
 PR 16-JAN-2001; 2001US-261938P.
 PR 16-JAN-2001; 2001US-262150P.
 PR 25-JAN-2001; 2001US-264395P.
 PR 02-FEB-2001; 2001US-266421P.
 PR 09-FEB-2001; 2001US-267623P.
 PR 28-FEB-2001; 2001WO-US06520.
 PR 09-MAR-2001; 2001US-274399P.
 PR 03-APR-2001; 2001US-280982P.
 PR 04-APR-2001; 2001US-282129P.
 PR 04-APR-2001; 2001US-282199P.
 PR 09-MAY-2001; 2001US-290589P.
 PR 25-MAY-2001; 2001WO-US17092.
 PR 01-JUN-2001; 2001WO-US17800.
 PR 20-JUN-2001; 2001WO-US19682.
 PR 29-JUN-2001; 2001WO-US21066.
 PR 09-JUL-2001; 2001WO-US21735.
 XX
 PA (GENTH) GENENTECH INC.
 XX
 PI Baker KP, Eaton DL, Filvaroff E, Goddard A, Grimaldi JC;
 PI Garney AL, Smith V, Stephan J, Watanabe CK, Wood WI, Zhang Z;
 PI Fong S;
 XX
 DR WPI; 2002-362426/39.
 DR N-PSDB; ABK69992.
 XX
 PT New PRO polypeptides and polynucleotides encoding the polypeptides,
 PT useful in gene therapy, chromosome identification, tissue typing, or
 PT for genetic analysis of individuals with genetic disorders -
 XX
 PS Claim 11; Figure 64; 218pp; English.
 XX
 CC This invention relates to the cDNA and protein sequences of novel
 CC secreted and transmembrane polypeptides PRO polypeptides. The
 CC invention also comprises a method for producing the proteins of the
 CC invention by recombinant means and antibodies specific for the protein
 CC of the invention. The antibody may be used for detecting the PRO
 CC proteins of the invention and may be used to modify their activity.
 CC polynucleotides may be used as hybridisation probes for a cDNA library
 CC to isolate the full-length PRO cDNA or to isolate other cDNAs, to
 CC construct hybridisation probes for mapping the gene which encodes that
 CC PRO and for genetic analysis of individuals with genetic disorders, in
 CC assays to identify other proteins or molecules involved in binding
 CC reaction, to generate transgenic animals or knock-out animals which in
 CC turn are useful in the development and screening of therapeutically
 CC useful reagents, for chromosome identification, and tissue typing. The

CC PRO polypeptides are useful in gene therapy, and as molecular weight
 CC markers for protein electrophoresis purposes. The sequences may
 CC also be used to detect overexpression on PRO polypeptides in cancerous
 CC tumours and for screening for differentially expressed genes using
 CC microarray technology. The present sequence represents a human PRO
 CC protein of the invention.
 CC
 XX
 SQ Sequence 213 AA;
 Query Match 100.0%; Score 1049; DB 23; Length 213;
 Best Local Similarity 100.0%; Pred. No. 1.7e-93;
 Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QGWAFFKNDATETIIPELGEYEPPELENNKTNRAENGRRPHHPEYKDVSSYSCREL 60
 DB 24 QGWAFFKNDATETIIPELGEYEPPELENNKTNRAENGRRPHHPEYKDVSSYSCREL 83
 QY 61 HFTRYVTGPGCRSAKPVTELVCSGCCGPARLLPNAIGRKKWRSSGDFRCIPRYRAQR 120
 DB 84 HFTRYVTGPGCRSAKPVTELVCSGCCGPARLLPNAIGRKKWRSSGDFRCIPRYRAQR 143
 QY 121 VOLLCPGGAPAPARKVRLVASCKCKRLTRFNOSLKDFTGARPPQGRKPRPARSAK 180
 DB 144 VOLLCPGGAPAPARKVRLVASCKCKRLTRFNOSLKDFTGARPPQGRKPRPARSAK 203
 QY 181 ANQAELENAY 190
 DB 204 ANQAELENAY 213
 XX
 RESULT 6
 ABB07209
 ID ABB07209 standard; Protein; 213 AA.
 XX
 AC ABB07209;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Human cloaked-2 polypeptide sequence.
 XX
 KW Cloaked-2; cysteine knot motif; nephrotropic; cardiact; immunomodulator;
 KW hepatotropic; antiinflammatory; antithyroid; cytostatic; neuroprotective;
 KW antianemic; hypotensive; antidiarrhythmic; antidiarrhythmoclerotic; muscular;
 KW antidiabetic; anorectic; gene therapy; cell therapy; antisense therapy;
 KW human.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..23
 FT /note= "signal peptide"
 FT Protein 24..213
 FT /note= "mature protein (ABB07207)"
 XX
 PN WO200192308-A2.
 XX
 PD 06-DEC-2001.
 XX
 PF 29-MAY-2001; 2001WO-US17478.
 XX
 PR 01-JUN-2000; 2000US-208550P.
 PR 04-AUG-2000; 2000US-223542P.
 XX
 PA (AMGE-) AMGEN INC.
 XX
 PI Paszty CJ, Gao Y;
 XX
 DR WPI; 2002-114325/15.
 DR N-PSDB; ABA94293.
 XX
 PT New human and mouse cysteine-knot polypeptide designated as Cloaked-2,
 PT for treating or preventing kidney, heart (e.g. myocardial infarction)
 PT or liver (e.g. hepatitis) diseases -

XX Example 1; Fig 1; 170pp; English.

CC The invention relates to polypeptides comprising a cysteine knot motif
CC and designated as Cloaked-2, derived from human and mouse. The cloaked-2
CC polypeptides can be expressed by standard recombinant methodology. The
CC cloaked-2 polynucleotides are useful in gene therapy and antisense
CC therapy. The cloaked-2 polypeptides and polynucleotides are useful for
CC treating, preventing, ameliorating or detecting diseases and disorders of
CC the kidney (e.g. anemia, hypertension or low blood pressure), heart (e.g.
CC cardiac hypertrophy, congestive heart failure, myocardial infarction,
CC arrhythmias, atherosclerosis, hypertension or low blood pressure),
CC skeletal muscle (e.g. muscular dystrophy or cachexia), placenta (e.g.
CC congenital abnormalities or miscarriage), liver (e.g. hepatitis (e.g.
CC cirrhosis), pancreas (e.g. diabetes or pancreatitis), thyroid (e.g.
CC Grave's disease or myxedema) or adrenal cortex (e.g. Cushing's disease
CC or Addison's disease), osteostasis or metabolic diseases (e.g. obesity,
CC cancer or myopathies), infections, or autoimmune diseases. Selective
CC binding agents may be used to modulate the biological activities of
CC cloaked-2 polypeptides or to detect cloaked-2 polypeptide levels in a
CC sample. Transgenic non-human animals are useful for drug candidate
CC screening. The present sequence represents the human cloaked-2
CC polypeptide sequence.

XX Sequence 213 AA;

Query Match 100.0%; Score 1049; DB 23; Length 213;
Best Local Similarity 100.0%; Pred. No. 1.7e-93;
Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGWAQKNDATETIIPELGEYEPPELENNKTNRRAENGSRPHHPFETKDVSEYSCREL 60
DB 24 QGWAQKNDATETIIPELGEYEPPELENNKTNRRAENGSRPHHPFETKDVSEYSCREL 83
QY 61 HFTRYVTDPGRSAKVPTELVCSGGCGPARLLPNAIGRKWMRPSGDPFRICIPRYRAQR 120
DB 84 HFTRYVTDPGRSAKVPTELVCSGGCGPARLLPNAIGRKWMRPSGDPFRICIPRYRAQR 143
QY 121 VOLLCRGEAPRARKVRLVASCCKRLTRFNQSELKDFTEARPOKGRPRPARSAK 180
DB 144 VOLLCRGEAPRARKVRLVASCCKRLTRFNQSELKDFTEARPOKGRPRPARSAK 203

QY 181 ANQAELENAY 190
DB 204 ANQAELENAY 213

RESULT 7
AAE17089
ID AAE17089 standard; Protein; 213 AA.

XX AAE17089;

XX 18-APR-2002 (first entry)

DE Human osteolevin protein.

XX Human; osteolevin; osteopathic; cytostratic; bone formation; osteoporosis;

KM Van Buchem-sclerosteosis disease; sclerosteosis; transgenic animal;

XX Paget's disease; chromosome 17.

XX Homo sapiens.

XX Key

XX Peptide

XX Protein

XX Misc-difference 10

XX /note= "During polymorphism wild type Val is substituted with Ile"

XX WO200198491-A2.

XX 27-DEC-2001.

XX 15-JUN-2001; 2001WO-EP06795.

XX 19-JUN-2000; 2000EP-0112867.

XX (HOFF) HOFMANN LA ROCHE & CO AG F.

XX (UNIV-) UNIV INTELLING ANTWERPEN UTA.

XX Balemans W, Ebelling M, Foernzler D, Patel N, Van Hul W,

XX Vickers BH;

XX WPI; 2002-139789/18.

XX N-PDB; AAD27576, AAD27577.

XX Claim 11; Fig 4; 70pp; English.

CC The invention relates to a nucleic acid encoding osteolevin region
CC polymorphisms. The invention also relates to genetic polymorphisms in
CC the Van Buchem-sclerosteosis disease region that are associated with
CC disorders resulting in either net excess bone formation or insufficient
CC bone formation in humans. Osteolevin DNA is useful for screening for
CC osteolevin polymorphisms associated with abnormal bone formation in a
CC subject and for the presence of a heritably linked form of abnormal bone
CC formation in a subject, by determining the presence of a polymorphism in
CC the osteolevin nucleic acid sequence obtained from the subject.
CC Osteolevin protein is useful for treating diseases associated with
CC abnormal bone formation, such as sclerosteosis, Van Buchem's disease and
CC Paget's disease. Nucleic acid which encode genes in the osteolevin
CC region or their modified forms can also be used to generate either
CC transgenic animals or knockout animals which are useful in the screening
CC and development of therapeutically useful reagents. Osteolevin proteins
CC are useful in pharmacological characterisation of novel modulators of the
CC activity of protein and protein complexes. Human osteolevin gene is
CC located on chromosome 17. The present sequence is human osteolevin
CC protein.

XX Sequence 213 AA;

Query Match 100.0%; Score 1049; DB 23; Length 213;
Best Local Similarity 100.0%; Pred. No. 1.7e-93;
Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGWAQKNDATETIIPELGEYEPPELENNKTNRRAENGSRPHHPFETKDVSEYSCREL 60
DB 24 QGWAQKNDATETIIPELGEYEPPELENNKTNRRAENGSRPHHPFETKDVSEYSCREL 83

QY 61 HFTRYVTDPGRSAKVPTELVCSGGCGPARLLPNAIGRKWMRPSGDPFRICIPRYRAQR 120

DB 84 HFTRYVTDPGRSAKVPTELVCSGGCGPARLLPNAIGRKWMRPSGDPFRICIPRYRAQR 143

QY 121 VOLLCRGEAPRARKVRLVASCCKRLTRFNQSELKDFTEARPOKGRPRPARSAK 180

DB 144 VOLLCRGEAPRARKVRLVASCCKRLTRFNQSELKDFTEARPOKGRPRPARSAK 203

QY 181 ANQAELENAY 190

DB 204 ANQAELENAY 213

RESULT 8

AAV96430
ID AAV96430 standard; Protein; 213 AA.

XX AAV96430;

XX 12-SEP-2000 (first entry)

DE Human TGF-beta binding protein (BEER) variant V101.
 XX osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
 KM BEER; variant; V101; gene therapy; antisense therapy; fracture;
 KM chromosome 17q12-21; bone mineralization.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 10 /label= V101
 FT /note="wild type valine has been substituted with
 FT isoleucine"
 XX
 FT MO200032773-A1.
 XX
 PD 08-JUN-2000.
 XX
 PD 24-NOV-1999; 99WO-US27990.
 XX
 PF 27-NOV-1998; 98US-0110283.
 XX
 PR (DARW-) DARWIN DISCOVERY LTD.
 PA
 PI Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepfer BW;
 PI Van Ness J, Winkler DG;
 XX
 DR N-PSDB; AAA29056.
 XX
 PT Nucleic acids (1) encoding a transforming growth factor beta binding
 PT protein, useful for identifying agents for treating osteopenia,
 PT osteoporosis and fractures
 XX
 PS Claim 3; Page 119-120; 162pp; English.
 XX
 CC This shows a variant human transforming growth factor-beta (TGF-beta)
 CC binding protein designated BEER V101, which comprises a substitution of
 CC isoleucine for the wild-type valine at residue 10. The CDNA and protein
 CC may be used for prevention, treatment and diagnosis of diseases
 CC associated with inappropriate BEER expression. For example, they may be
 CC used to treat disorders associated with decreased TGF-beta BP expression.
 CC The CDNA or vectors may be administered to treat diseases by rectifying
 CC mutations or deletions in a patient's genome that affect the activity of
 CC BEER by expressing inactive proteins or to supplement the patients own
 CC production of BEER polypeptides. The nucleic acids may be used for
 CC recombinant production of BEER, gene therapy, antisense therapy, as
 CC probes for diagnostic assays and for functional studies. BEER may be used
 CC to raise antibodies and for identification of BEER modulators. BEER
 CC antagonists may be used to increase bone mineral content for the
 CC treatment of disorders such as osteopenia, osteoporosis, fractures and
 CC other disorders associated with low mineral content.
 XX
 SQ Sequence 213 AA;
 XX
 Query Match 99.1%; Score 1040; DB 21; Length 213;
 Best Local Similarity 99.5%; Pred. No. 1.3e-92;
 Matches 189; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 XX
 QY 1 QGWAQKNDATETIPELGEYEPPEPELENNKTNRAENGGRPPHPPETKDVSEYSGREL 60
 DB 24 QGWAQKNDATETIPELGEYEPPEPELENNKTNRAENGGRPPHPPETKDVSEYSGREL 83
 QY 61 HFTTYVTDGPGRSAPPTVELVCSGCCGPARLLPVAIGKMMWRSPGDFPCIDRYYAQR 120
 DB 84 HFTTYVTDGPGRSAPPTVELVCSGCCGPARLLPVAIGKMMWRSPGDFPCIDRYYAQR 143
 QY 121 VOLLCPGGEAPRARKVTLVASCCKRLTRPHNOSELKDFGTEAPRPOKGRPPRARSASAK 180
 DB 144 VOLLCPGGEAPRARKVTLVASCCKRLTRPHNOSELKDFGTEAPRPOKGRPPRARSASAK 203
 QY 181 ANQAELENAY 190
 |||||||

DB 204 ANQAELENAY 213
 RESULT 9
 ID AAY96436 standard; Protein; 213 AA.
 XX
 AC AAY96436;
 XX
 DT 12-SEP-2000 (first entry)
 XX
 DE Human TGF-beta binding protein (BEER) variant P38R.
 XX
 KM osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
 KM BEER; variant; P38R; gene therapy; antisense therapy; fracture;
 KM chromosome 17q12-21; bone mineralization.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 38 /label= P38R
 FT /note="wild type proline has been substituted with
 FT arginine"
 XX
 FT MO200032773-A1.
 XX
 PD 08-JUN-2000.
 XX
 PD 24-NOV-1999; 99WO-US27990.
 XX
 PF 27-NOV-1998; 98US-0110283.
 XX
 PR (DARW-) DARWIN DISCOVERY LTD.
 PA
 PI Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepfer BW;
 PI Van Ness J, Winkler DG;
 XX
 DR N-PSDB; AAA29062.
 XX
 PT Nucleic acids (1) encoding a transforming growth factor beta binding
 PT protein, useful for identifying agents for treating osteopenia,
 PT osteoporosis and fractures
 XX
 PS Disclosure; Page 121; 162pp; English.
 XX
 CC This shows a variant human transforming growth factor-beta
 CC (TGF-beta) binding protein designated BEER P38R. The encoded protein
 CC comprises a substitution of arginine for the wild-type proline at
 CC residue 38. The CDNA and protein may be used for prevention, treatment
 CC and diagnosis of diseases associated with inappropriate BEER expression.
 CC For example, they may be used to treat disorders associated with
 CC decreased TGF-beta BP expression. The CDNA or vectors may be administered
 CC to treat diseases by rectifying mutations or deletions in a patient's
 CC genome that affect the activity of BEER by expressing inactive proteins
 CC or to supplement the patients own production of BEER polypeptides. The
 CC nucleic acids may be used for recombinant production of BEER, gene
 CC therapy, antisense therapy, as probes for diagnostic assays and for
 CC functional studies. BEER may be used to raise antibodies and for
 CC identification of BEER modulators. BEER antagonists may be used to
 CC increase bone mineral content for the treatment of disorders such as
 CC osteopenia, osteoporosis, fractures and other disorders associated with
 CC low mineral content.
 XX
 SQ Sequence 213 AA;
 XX
 Query Match 99.1%; Score 1040; DB 21; Length 213;
 Best Local Similarity 99.5%; Pred. No. 1.3e-92;
 Matches 189; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 XX
 QY 1 QGWAQKNDATETIPELGEYEPPEPELENNKTNRAENGGRPPHPPETKDVSEYSGREL 60
 |||||||

DB 24 OGWAFKNDATETIIRLGEYPEPPELENNKTMNRAENGRRPHHPFTKDVSEYSCREL 83
 QY 61 HFTRYVTDGCRSAKPYTELVCSGCGPARRLLPNAIGRGKWMRPSGDPFRCTPDRYRQR 120
 DB 84 HFTRYVTDGCRSAKPYTELVCSGCGPARRLLPNAIGRGKWMRPSGDPFRCTPDRYRQR 143
 QY 121 VOLLCRGEAPARRKRVRLVASCCKRLTRFNQSELKDFGTEAARPOKGRPRPARSAK 180
 DB 144 VOLLCRGEAPARRKRVRLVASCCKRLTRFNQSELKDFGTEAARPOKGRPRPARSAK 203
 QY 181 ANQAELENNAY 190
 DB 204 ANQAELENNAY 213

RESULT 10
 ID AAY96431 standard; Protein; 213 AA.
 AAY96431;

12-SEP-2000 (first entry)
 Vervet TGF-beta binding protein (BEER).
 osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
 BEER; gene therapy; antisense therapy; fracture; bone mineralization.
 Cercopithecus pygerythrus.

WO200032773-A1.

08-JUN-2000.

24-NOV-1999; 99WO-US27990.

27-NOV-1998; 98US-0110283.

(DARW-) DARWIN DISCOVERY LTD.

Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepfer BW;
 Van Ness J, Winkler DG;

WPI; 2000-412321/35.
 N-PSDB; AAA29057.

Nucleic acids (1) encoding a transforming growth factor beta binding
 protein, useful for identifying agents for treating osteopenia,
 osteoporosis and fractures

Claim 4; Page 122-123; 162pp; English.

This shows a vervet transforming growth factor-beta (TGF-beta) binding
 binding protein designated VBER. The cDNA and protein may be used for
 prevention, treatment and diagnosis of diseases associated with
 inappropriate BEER expression. For example, they may be used to treat
 disorders associated with decreased TGF-beta BP expression. The cDNA or
 vectors may be administered to treat diseases by rectifying mutations or
 deletions in a patient's genome that affect the activity of BEER by
 expressing inactive proteins or to supplement the patient's own production
 of BEER polypeptides. The nucleic acids may be used for recombinant
 production of BEER, gene therapy, antisense therapy, as probes for
 diagnostic assays and for functional studies. BEER may be used to raise
 antibodies and for identification of BEER modulators. BEER antagonists
 may be used to increase bone mineral content for the treatment of
 disorders such as osteopenia, osteoporosis, fractures and other disorders
 associated with low mineral content.

Sequence 213 AA;

Query Match 98.5%; Score 1033; DB 21; Length 213;
 Best Local Similarity 98.4%; Pred. No. 6.2e-92;
 Matches 187; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 OGWAFKNDATETIIRLGEYPEPPELENNKTMNRAENGRRPHHPFTKDVSEYSCREL 60
 DB 24 OGWAFKNDATETIIRLGEYPEPPELENNKTMNRAENGRRPHHPFTKDVSEYSCREL 83
 QY 61 HFTRYVTDGCRSAKPYTELVCSGCGPARRLLPNAIGRGKWMRPSGDPFRCTPDRYRQR 120
 DB 84 HFTRYVTDGCRSAKPYTELVCSGCGPARRLLPNAIGRGKWMRPSGDPFRCTPDRYRQR 143
 QY 121 VOLLCRGEAPARRKRVRLVASCCKRLTRFNQSELKDFGTEAARPOKGRPRPARSAK 180
 DB 144 VOLLCRGEAPARRKRVRLVASCCKRLTRFNQSELKDFGTEAARPOKGRPRPARSAK 203
 QY 181 ANQAELENNAY 190
 DB 204 ANQAELENNAY 213

RESULT 11
 ID AAY96433 standard; Protein; 213 AA.
 AAY96433;

12-SEP-2000 (first entry)
 Rat TGF-beta binding protein (BEER).

osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
 BEER; gene therapy; antisense therapy; fracture; bone mineralization.

Rattus norvegicus.

WO200032773-A1.

08-JUN-2000.

24-NOV-1999; 99WO-US27990.

27-NOV-1998; 98US-0110283.

(DARW-) DARWIN DISCOVERY LTD.

Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepfer BW;
 Van Ness J, Winkler DG;

WPI; 2000-412321/35.
 N-PSDB; AAA29059.

Nucleic acids (1) encoding a transforming growth factor beta binding
 protein, useful for identifying agents for treating osteopenia,
 osteoporosis and fractures

Claim 6; Page 125-126; 162pp; English.

This shows a rat transforming growth factor-beta (TGF-beta) binding
 protein designated RBER. The cDNA and protein may be used for
 prevention, treatment and diagnosis of diseases associated with
 inappropriate BEER expression. For example, they may be used to treat
 disorders associated with decreased TGF-beta BP expression. The cDNA or
 vectors may be administered to treat diseases by rectifying mutations or
 deletions in a patient's genome that affect the activity of BEER by
 expressing inactive proteins or to supplement the patient's own production
 of BEER polypeptides. The nucleic acids may be used for recombinant
 production of BEER, gene therapy, antisense therapy, as probes for
 diagnostic assays and for functional studies. BEER may be used to raise
 antibodies and for identification of BEER modulators. BEER antagonists
 may be used to increase bone mineral content for the treatment of
 disorders such as osteopenia, osteoporosis, fractures and other disorders
 associated with low mineral content.

Sequence 213 AA;

Query Match 92.9%; Score 974; DB 21; Length 213;
 Best Local Similarity 92.1%; Pred. No. 3.2e-86;
 Matches 175; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 QGWAFFKNDATETIIPELGEYEPPEPELENNKTNNRAENGRRPHHPETKDVSEYSCREL 60
 DB 24 QGWAFFKNDATETIIPELGEYEPPEPELENNKTNNRAENGRRPHHPETKDVSEYSCREL 83
 QY 61 HPTVYVTDGPRSAKPVTELVCSSGCCPARLLPNAIGRKWMPSPGDPFCIDRRYAQR 120
 DB 84 HPTVYVTDGPRSAKPVTELVCSSGCCPARLLPNAIGRKWMPSPGDPFCIDRRYAQR 143
 QY 121 VOLLCPGGAAPRAKRVLVASCKCKRLTRFNQSELKDPGTEARPOGKRKPRPARSAK 180
 DB 144 VOLLCPGGAAPRAKRVLVASCKCKRLTRFNQSELKDPGTEARPOGKRKPRPARSAK 203
 QY 181 ANQAELENAY 190
 DB 204 ANQAELENAY 213

RESULT 12
 AAB26105
 ID AAB26105 standard; Protein; 367 AA.
 AC AAB26105;
 XX
 DT 15-JAN-2001 (first entry)
 DE Human DAN/Cerberus-related protein 6 (hDCR6) #1.
 XX
 KM Human; DNA/Cerberus-related protein 6; hDCR6; morphogenic protein;
 KM antagonist; BMP; cell growth; cell differentiation; bone formation;
 KM gene therapy.
 OS Homo sapiens.
 XX
 PN MO200055193-A2.
 XX
 PD 21-SEP-2000.
 XX
 PF 02-MAR-2000; 2000WO-US05537.
 XX
 PR 12-MAR-1999; 99US-0124118.
 XX
 PA (REGG-) REGENERON PHARM INC.
 XX
 PI Economides AN;
 XX
 DR WPI; 2000-638179/61.
 DR N-PSDB; AAA94049, AAA94050.
 XX
 PT Novel isolated, human DNA/Cerberus related protein 6 which include
 PT natural homologue, and polypeptides comprising DCR6 domain and nucleic
 PT acids encoding the proteins which are useful as probes and primers -
 XX
 PS Claim 7; Fig 2; 40pp; English.
 XX
 CC The present sequence comprises the human DAN/Cerberus-related protein 6
 CC (hDCR6). Its coding sequence was isolated from a genomic DNA clone
 CC following identification using computer-based "virtual cloning". hDCR6
 CC is closely related to the DAN and DCR5 proteins, both of which act as
 CC antagonists of morphogenic proteins such as BMP. It is possible that
 CC the hDCR6 gene and protein can be used as immunogens, modulators of cell
 CC function, growth and differentiation, to reduce undesirable bone
 CC formation, to identify DCR6 binding agents, in diagnosis, and in gene
 CC therapy.
 CC
 SQ Sequence 367 AA;

Query Match 90.9%; Score 954; DB 21; Length 367;
 Best Local Similarity 54.9%; Pred. No. 5.2e-84;
 Matches 189; Conservative 0; Mismatches 1; Indels 154; Gaps 1;

QY 1 QGWAFFKNDATETIIPELGEYEPPEPELENNKTNNRAENGRRPHHPETKD----- 51
 DB 24 QGWAFFKNDATETIIPELGEYEPPEPELENNKTNNRAENGRRPHHPETALVVEQGLE 83
 QY 52 ----- 51
 DB 84 VGEWLAGAAAFKRCRGRNROTLVRAQREDAGVVRVHOGIRTSQAQKKKGFESPPG 143
 QY 52 ----- 51
 DB 144 NIGATSSCWTHWEGNKVAEKSTAQAPQPPPPNNLLGLWAKKGCNCGGPPSVRME 203
 QY 52 -----VSEYSCRELHPTVYVTDGPRSAKPVTELVCSSGCC 86
 DB 204 DKGASPHSPSPNQGLGSSDCLKGVSEYSCRELHPTVYVTDGPRSAKPVTELVCSSGCC 263
 QY 87 GPARLLPNAIGRKWMPSPGDPFCIDRRYAQRVOLLCPGGAAPRAKRVLVASCKCKR 146
 DB 264 GPARLLPNAIGRKWMPSPGDPFCIDRRYAQRVOLLCPGGAAPRAKRVLVASCKCKR 323
 QY 147 LTRFNQSELKDPGTEARPOGKRKPRPARSAKANAQAELENAY 190
 DB 324 LTRFNQSELKDPGTEARPOGKRKPRPARSAKANAQAELENAY 367

RESULT 13
 ABB07208
 ID ABB07208 standard; Protein; 188 AA.
 AC ABB07208;
 XX
 DT 26-MAR-2002 (first entry)
 DE Mouse cloaked-2 polypeptide mature protein sequence.
 XX
 KM Cloaked-2; cysteine knot motif; nephrotropic; cardiant; immunomodulator;
 KM hepatocrotic; antiinflammatory; antithyroid; cytosolic; neuroprotective;
 KM antianemic; hypotensive; antiarrhythmic; antiarteriosclerotic; muscular;
 KM antidiabetic; anorectic; gene therapy; cell therapy; antisense therapy;
 KM mouse.
 XX
 OS Mus musculus.
 XX
 FH Key location/Qualifiers
 FT Misc-difference 175..177
 FT /note="this region is missing in the sequence provided
 FT in the sequence listing but has been indicated
 FT correctly in the sequence in the Figure"
 XX
 PN WO200192308-A2.
 XX
 PD 06-DEC-2001.
 XX
 PF 29-MAY-2001; 2001WO-US17478.
 XX
 PR 01-JUN-2000; 2000US-208550P.
 PR 04-AUG-2000; 2000US-223542P.
 XX
 PA (AMGE-) AMGEN INC.
 XX
 PI Paszty CJ, Gao Y;
 XX
 DR WPI; 2002-114325/15.
 DR N-PSDB; ABA94294.
 XX
 PT New human and mouse cysteine-knot polypeptide designated as Cloaked-2,
 PT for treating or preventing kidney, heart (e.g. myocardial infarction)
 PT or liver (e.g. hepatitis) diseases -
 XX
 PS Claim 13; Fig 2; 170pp; English.
 XX
 CC The invention relates to polypeptides comprising a cysteine knot motif

Oy	723 ACCCGGGCCCACTTTCGTCTCCTGCGCGTGATTT	759
Db	748 ACCCGGGCCCACTTTCGTCTCCTGCGCGTGATTT	784

RESULT 6

ID ABK69992 standard; DNA; 2329 BP.

AC ABK69992;

DT 15-JUL-2002 (first entry)

DE CDNA encoding human Pro peptide #32.

KM Human; ss; gene; PRO; secreted protein; transmembrane protein;
KM genetic disorder; tumour; cancer.

OS Homo sapiens.

PN W0200224888-A2.

PD 28-MAR-2002.

PF 29-AUG-2001; 2001WO-US27099.

PR 01-SEP-2000; 2000US-229896P.

PR 22-SEP-2000; 2000US-235147P.

PR 12-JAN-2001; 2001US-261878P.

PR 16-JAN-2001; 2001US-261939P.

PR 25-JAN-2001; 2001US-264395P.

PR 09-FEB-2001; 2001US-267623P.

PR 09-MAR-2001; 2001US-274399P.

PR 04-APR-2001; 2001US-282129P.

PR 09-MAY-2001; 2001US-290589P.

PR 01-JUN-2001; 2001WO-US17800

PR 29-JUN-2001; 2001WO-US21066

XX
XX
CENY) CENYETECU TWO

2000

PI Gurney AL, Smith V, Stephen

XX 363436/300
WDT -

DR P-PSDB; ABG34061.
XX

PT New PRO polypeptides and polypeptidases

PT for genetic analysis of indi-
 viduals

PS Claim 2; Figure 63; 218pp; 1

CC This invention relates to the

CC invention also comprises

of the invention. The antibody may be used for detecting the

polynucleotides may be used as hybridisation probes for a cDNA library.

CC	construct hybridisation probes for mapping the gene which encodes that
CC	PRO and for genetic analysis of individuals with genetic disorders. In
CC	assays to identify other proteins or molecules involved in binding
CC	reaction, to generate transgenic animals or knock-out animals which in
CC	turn are useful in the development and screening of therapeutically
CC	useful reagents, for chromosome identification, and tissue typing. The
CC	PRO polypeptides are useful in gene therapy, and as molecular weight
CC	markers for protein electrophoresis purposes. The sequences may
CC	also be used to detect overexpression on PRO polypeptides in cancerous
CC	tumours and for screening for differentially expressed genes using
CC	microarray technology. The present sequence represents a cDNA encoding
CC	a human PRO protein of the invention.
XX	
XX	Sequence 2329 BP; 587 A; 645 C; 612 G; 485 T; 0 other;
XX	
XX	Query Match 99.7%; Score 757; DB 24; Length 2329;
XX	Best Local Similarity 100.0%; Pred. No. 1.1e-137;
XX	Matches 757; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	3 CTGGAAGTGGCGTGCCTCTCTGTGCTGTGACCATGACAGCTCCCACTGCGCTGTGTCT 62
DB	28 CTGGAAGTGGCGTGCCTCTCTGTGCTGTGACCATGACAGCTCCCACTGCGCTGTGTCT 87
QY	63 CGTCTGCTGTGTGACACAGAGCTTCCGTGTGTGTGAGAGGCCAGAGGTGTGGACGCGTT 122
DB	88 CGTCTGCTGTGTGACACAGAGCTTCCGTGTGTGTGAGAGGCCAGAGGTGTGGACGCGTT 147
QY	123 CAAGAATGATGCACAGGAATTCATCCCGAGCTCGAAGATACCCCGAGCTTCACCGGA 182
DB	148 CAGAATATATGCCAGGAATTCATCCCGAGCTCGAAGATACCCCGAGCTTCACCGGA 207
QY	183 GCTGGAGAACAACAAGACCATGAACCGGCGGAGAACCGAGGGCGGCTCCCAACACC 242
DB	208 GCTGGAGAACAACAAGACCATGAACCGGCGGAGAACCGAGGGCGGCTCCCAACACC 267
QY	243 CTTTGAGACCAAGAAGCTGTCCGAGTACAGCTGCCGCGAGCTGCATTCACCCGCTTACT 302
DB	268 CTTTGAGACCAAGAAGCTGTCCGAGTACAGCTGCCGCGAGCTGCATTCACCCGCTTACT 327
QY	303 GACCGATGGGCGCGTGCAGAGCGCAACCGGTCACCGAGCTGGTGTCTCCGCGAGTGG 362
DB	328 GACCGATGGGCGCGTGCAGAGCGCAACCGGTCACCGAGCTGGTGTCTCCGCGAGTGG 387
QY	363 CGGCGCGGCGCGCTGTGTGCCCAAGCCATCGGCGCGCGCAAGTGTGGCGCACTTATGG 422
DB	388 CGGCGCGGCGCGCTGTGTGCCCAAGCCATCGGCGCGCGCAAGTGTGGCGCACTTATGG 447
QY	423 GCCCGACTTCCGCTGCATCCCGAACCGGTACCGCGCGCGAGCGCGTGCAGCTGTGTCTC 482
DB	448 GCCCGACTTCCGCTGCATCCCGAACCGGTACCGCGCGCGAGCGCGTGCAGCTGTGTCTC 507
QY	483 CGGTGTGTAGAGGCGCGCGCGCGCGCGGAAGTGTGCTGTGTGCTCTGTGTCAAGTGAACG 542
DB	508 CGGTGTGTAGAGGCGCGCGCGCGCGCGGAAGTGTGCTGTGTGCTCTGTGTCAAGTGAACG 567
QY	543 CTTCAACCGGCTTCCCAACACAGTGCAGACTCAAGGACTTTCGGAGCCGAGGCGGCTCGGCC 602
DB	568 CTTCAACCGGCTTCCCAACACAGTGCAGACTCAAGGACTTTCGGAGCCGAGGCGGCTCGGCC 627
QY	603 GCAGAGGAGCGGAGAGCCGCGGCGCTCCGCGCGCGAGCGCCAAAGCCAAACAGGCGGACT 662
DB	628 GCAGAGGAGCGGAGAGCCGCGGCGCTCCGCGCGCGAGCGCCAAAGCCAAACAGGCGGACT 687
QY	663 GAGAGAGCGCTTACTAGAGCCGCGCGCGCTCTTCCACCGAGCGGAGCGCCCGGCTCTGA 722

RESULT 7

ID	AAA29056
XX	AAA29056 standard; cDNA; 2301 BP.
AC	AAA29056;
XX	
DT	12-SEP-2000 (first entry)
XX	
DE	Human TGF-beta binding protein (BEER) variant V101 cDNA.
KW	osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
KM	BEER; Variant; V101; gene therapy; antisense therapy; fracture;
XX	Chromosome 17q12-21; bone mineralization; ss.
OS	Homo sapiens.
XX	
FH	Key
FT	Location/Qualifiers
CDS	48..689
FT	/tag= a
FT	/label= BEER variant_V101
FT	/product= TGF-beta_binding_protein
XX	
PN	WO200032773-A1.
PD	
XX	08-JUN-2000.
Pf	
XX	24-NOV-1999; 99WO-US27990.
PR	
XX	27-NOV-1998; 98US-0110283.
XX	
PA	(DARW-) DARWIN DISCOVERY LTD.
PI	Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepfer BW;
PT	Van Ness J, Winkler DG;
XX	
DR	WPI: 2000-412321/35.
DR	P-PSDB; AAY96430.
XX	
PT	Nucleic acids (1) encoding a transforming growth factor beta binding
PT	protein, useful for identifying agents for treating osteopenia,
PT	osteoporosis and fractures
PS	Claim 1; Page 118-119; 162pp; English.
CC	This cDNA encodes a variant human transforming growth factor-beta
CC	(TGF-beta) binding protein designated BEER V101. The encoded protein
CC	comprises a substitution of isoleucine for the wild-type valine at
CC	residue 10. The cDNA and protein may be used for prevention, treatment
CC	and diagnosis of diseases associated with inappropriate BEER expression.
CC	For example, they may be used to treat disorders associated with
CC	decreased TGF-beta BP expression. The cDNA or vectors may be administered
CC	to treat diseases by rectifying mutations or deletions in a patient's
CC	genome that affect the activity of BEER by expressing inactive proteins
CC	or to supplement the patients own production of BEER polypeptides. The
CC	nucleic acids may be used for recombinant production of BEER, gene
CC	therapy, antisense therapy, as probes for diagnostic assays and for
CC	functional analysis. BEER may be used to raise antibodies and for
CC	identification of BEER modulators. BEER antagonists may be used to
CC	increase bone mineral content for the treatment of disorders such as
CC	osteopenia, osteoporosis, fractures and other disorders associated with
CC	low mineral content.
CC	
SO	Sequence 2301 BP; 569 A; 634 C; 614 G; 484 T; 0 other;
Query Match	99.6%; Score 755.8; DB 21; Length 2301;
Best Local Similarity	99.7%; Fred. No. 1.8e-13;
Matches 757; Conservative	0; Mismatches 2; Indels 0; Gaps 0;
QY	1 TACTGGAAGGTGGCGCTTCCTCTGCGTGATGCATGCACTGCCAATGGCCCTGTGT 60
Dd	12 TACTGGAAGGTGGCGCTTCCTCTGCGTGATGCATGCACTGCCAATGGCCCTGTGT 71
OY	61 CTCGCTGCGCTGCTGTATCACACAGCCTTCGCTGTATGTGAAGGCCCAAGTGGCGGCG 120

[illegible]

XX	MO200198491-A2.
PD	27-DEC-2001.
PF	15-JUN-2001; 200IWO-EP06795.
PR	19-JUN-2000; 2000EP-0112867.
XX	(HOFF) HOFFMANN LA ROCHE & CO AG F.
PA	(UYIN-) UNIV INSTELLING ANTWERPEN UIA.
XX	Balemans W, Ebeling M, Poernzler D, Patel N, Van Hul W;
P1	Vickery BH;
DR	WPI: 2002-139789/18.
XX	P-PsDB; AAE17089.
PT	Novel genetic polymorphisms in the Van Buchem-sclerosteosis disease
FT	region that are associated with abnormal bone formation useful for
PS	diagnosis and assessment of osteoporosis or sclerosteosis in humans
XX	-
PS	Claim 10; Fig 1B; 70pp; English.
CC	The invention relates to a nucleic acid encoding osteolevin region
CC	polymorphisms. The invention also relates to genetic polymorphisms in
CC	the Van Buchem-sclerosteosis disease region that are associated with
CC	disorders resulting in either net excess bone formation or insufficient
CC	bone formation in humans. Osteolevin DNA is useful for screening for
CC	osteolevin polymorphisms associated with abnormal bone formation in a
CC	subject and for the presence of a heritably linked form of abnormal bone
CC	formation in a subject, by determining the presence of a polymorphism in
CC	the osteolevin nucleic acid sequence obtained from the subject.
CC	Osteolevin protein is useful for treating diseases associated with
CC	abnormal bone formation, such as sclerosteosis, Van Buchem's disease and
CC	Paget's disease. Nucleic acids which encode genes in the osteolevin
CC	region or their modified forms can also be used to generate either
CC	transgenic animals or knockout animals which are useful in the screening
CC	and development of therapeutically useful reagents. Osteolevin proteins
CC	are useful in pharmacological characterisation of novel modulators of the
CC	activity of protein and protein complexes. Human osteolevin gene is
CC	located on chromosome 17. The present sequence is human osteolevin cDNA.
XX	
SQ	Sequence 2271 BP; 573 A; 623 C; 599 G; 476 T; 0 other;
	Query Match 96.8%; Score 735; DB 24; Length 2271;
	Best Local Similarity 100.0%; Pred. No. 1.9e-133;
	Matches 735; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY	25 CTGGCTGTGATCATGAGTCCCACTGGCGCCTGTGTCTGTCTGCCTGTGTACACA 84
Db	1 CTGGCTGTGATCATGAGTCCCACTGGCGCCTGTGTCTGTCTGTGTGTGATCACACA 60
OY	85 GCCTTCGGTGATGTGAAGAGGCCAGGGGTGGAGAGCCTTCAAGAATGATGCCAGAAATC 144
Db	61 GCCTTCGGTGATGTGAAGAGGCCAGGGGTGGAGAGCCTTCAAGAATATATGCCAGAAATC 120
OY	145 ATCCCCGAGCTCGAGAGTAGTACCCTGAGCTTCCAACGAGCTGTGAGAACAAACAAGACCATG 204
Db	121 ATCCCCGAGCTCGAGAGTAGTACCCTGAGCTTCCAACGAGCTGTGAGAACAAACAAGACCATG 180
OY	205 AACCGGGCGGAGAAACGAGAGGGGGGCTCTCCCAACACACCCCTTTGAGACCAAAGCTGTGTC 264
Db	181 AACCGGGCGGAGAAACGAGAGGGGGGCTCTCCCAACACACCCCTTTGAGACCAAAGCTGTGTC 240
OY	265 GAGTACAGCTGCGCGAGAGCTGACCTTACCCCGCTTAAGTACGATGAGGCGGTGCGCAGC 324
Db	241 GAGTACAGCTGCGCGAGAGCTGACCTTACCCCGCTTAAGTACGATGAGGCGGTGCGCAGC 300
OY	325 GCCAAGCCGGTCAACCGAGCTGTGTGCTCCGGCCAGTGTGGCCCGCGCGGCTGTGCTGCC 384
Db	301 GCCAAGCCGGTCAACCGAGCTGTGTGCTCCGGCCAGTGTGGGCCCGCGCGGCTGTGCTGCC 360
OY	385 AACGCCATGCGCGCGCAAGTGTGTGCGACCTTAGTGTGGCCGCACTTCCGCTGCATCCCC 444

Db	361	AAAGCCATCGGCGCGGCAAGTGGTGGGACCTTAGTGAGGCCCGACTTCCGCTGCATCCCC	420
Oy	445	GACCGCTACCGCGCGCGAGCGCTGAGACTGCTGTGTCCCGTGTGAGGCGCCGCGCG	504
Db	421	GACCGCTACCGCGCGCGAGCGCTGCACTGTGTGTCCCGTGTGAGGCGCGCGCG	480
Oy	505	CGCAAGGTGGCGCTGTGTGCGCTCGTGCAGTGCAGAGGCGCTCACCCGCTTCCACAACGAG	564
Db	481	CGCAAGGTGGCGCTGTGTGCGCTCGTGCAGTGCAGAGGCGCTCACCCGCTTCCACAACGAG	540
Oy	565	TCGAGACTCAAGGACTTCGAGACCGAGGCGCGCTCGCGCGAGAGGAGCGGACCGCG	624
Db	541	TCGAGACTCAAGGACTTCGAGACCGAGGCGCGCTCGCGCGAGAGGAGCGGACCGCG	600
Oy	625	CCCCGCGCCCGGAGCGCCAAAGCCAAAGCGAGCTGAGAAAGCTTACAGAGCCCG	684
Db	601	CCCCGCGCCCGGAGCGCCAAAGCCAAAGCGAGCTGAGAAAGCTTACAGAGCCCG	660
Oy	685	CCGCGCGCCCTTCCCCACCGCGGCGGCGCCCGCGCTTGAACCCCGCCCACTTCTGTC	744
Db	661	CCGCGCGCCCTTCCCCACCGCGGCGGCGCCCGCGCTTGAACCCCGCCCACTTCTGTC	720
Oy	745	CTCTGCGCGTGTGT 759	
Db	721	CTCTGCGCGTGTGT 735	
RESULT 9			
AAA94051			
ID	AAA94051	standard; DNA; 642 BP.	
AC	AAA94051;		
XX			
DT	15-JAN-2001	(first entry)	
XX			
DE	Human DAN/Cerberus-related protein 6 (hDcR6) cDNA exons 1 and 4.		
XX			
KW	Human; DNA/Cerberus-related protein 6; hDcR6; morphogenic protein;		
KW	antagonist; BMP; cell growth; cell differentiation; bone formation;		
XX	gene therapy; ss.		
XX			
OS	Homo sapiens.		
XX			
Key	Location/Qualifiers		
XX			
FT	CDS	1..642	
FT		/*tag= a	
FT		/product= "hDcR6 #2"	
FT		/partial	
XX			
XX	WO200055193-A2.		
XX			
XX	21-SEP-2000.		
XX			
XX	02-MAR-2000; 2000WO-US05537.		
XX			
XX	12-MAR-1999; 99US-0124118.		
XX			
PA	(REGC-) REGENERON PHARM INC.		
XX			
XX	Economides AN;		
XX			
XX	WPI; 2000-638179/61.		
XX			
XX	P-PSDB; AAB26106.		
XX			
PT	Novel isolated, human DNA/Cerberus related protein 6 which include		
PT	natural homologue, and polypeptides comprising DCR6 domain and nucleic		
PT	acids encoding the proteins which are useful as probes and primers -		
XX			
XX	Claim 2; Fig 3; 40pp; English.		
XX			
CC	The present sequence comprises exons 1 and 4 of the human		
CC	DAN/Cerberus-related protein 6 (hDcR6) coding sequence. It was isolated		

QY 397 CGCGCAAGTGTGGGACGCTAGTGGGCGGACTTCGCTGTCATCCCGGACGCTACCG 456
 Db 361 CGCGCAAGTGTGGGACGCTAGTGGGCGGACTTCGCTGTCATCCCGGACGCTACCG 420
 QY 457 GCGGAGCGGTGACGCTGTGTGTCCCGGTGTGTGAGGCGCGCGCGGCAAGGTGCGC 516
 Db 421 GCGGAGCGGTGACGCTGTGTGTCCCGGTGTGTGAGGCGCGCGCGGCAAGGTGCGC 480
 QY 517 CTGTGTGCTTCGTGACGCTGACGCTGACGCTGACGCTGACGCTGACGCTGACGCT 576
 Db 481 CTGTGTGCTTCGTGACGCTGACGCTGACGCTGACGCTGACGCTGACGCTGACGCT 540
 QY 577 GACTTCGCGACCGGAGCGCGCTCGGCGCGGACGCGGCGGCGGCGGCGGCGGCGG 636
 Db 541 GACTTCGCGTCCGAGGCGCGCTCGGCGCGGACGCGGCGGCGGCGGCGGCGGCGG 600
 QY 637 AGCGCCAAAGCCAAACGAGCGCGAGCTTGAGAGACGCTACTAG 678
 Db 601 GGGGCAAGCCAAATCAGCGCGAGCTGAGAGACGCTACTAG 642
 RESULT 11
 ID AAA29059 standard; cDNA; 674 BP.
 XX AAA29059;
 AC AAA29059;
 DT 12-SEP-2000 (first entry)
 XX
 DE Rat TGF-beta binding protein (BBER) cDNA.
 XX
 KM osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
 KM BBER; gene therapy; antisense therapy; fracture; bone mineralization; ss.
 XX
 OS Rattus norvegicus.
 XX
 FT Key Location/Qualifiers
 FT CDS 33..674
 FT /tag= a
 FT /product= TGF-beta_binding_protein
 FT
 XX MO200032773-A1.
 PN 08-JUN-2000.
 XX
 PF 24-NOV-1999; 99WO-US27990.
 XX
 PR 27-NOV-1998; 98US-0110283.
 XX
 PA (DARW-) DARWIN DISCOVERY LTD.
 XX
 PI Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepker BW;
 PI Van Ness J, Winkler DG;
 XX
 DR WPI; 2000-412321/35.
 DR P-PSDB; AAY96433.
 XX
 PT Nucleic acids (1) encoding a transforming growth factor beta binding
 PT protein, useful for identifying agents for treating osteopenia,
 PT osteoporosis and fractures
 XX
 PS Claim 1; Page 125; 162pp; English.
 XX
 CC This cDNA encodes a rat transforming growth factor-beta (TGF-beta)
 CC binding protein designated BBER. The cDNA and protein may be used for
 CC prevention, treatment and diagnosis of diseases associated with
 CC inappropriate BBER expression. For example, they may be used to treat
 CC disorders associated with decreased TGF-beta BP expression. The cDNA or
 CC vectors may be administered to treat diseases by rectifying mutations or
 CC deletions in a patient's genome that affect the activity of BBER by
 CC expressing inactive proteins or to supplement the patients own production
 CC of BBER polypeptides. The nucleic acids may be used for recombinant

CC production of BBER, gene therapy, antisense therapy, as probes for
 CC diagnostic assays and for functional studies. BBER may be used to raise
 CC antibodies and for identification of BBER modulators. BBER antagonists
 CC may be used to increase bone mineral content for the treatment of
 CC disorders such as osteopenia, osteoporosis, fractures and other disorders
 CC associated with low mineral content.
 XX
 SQ Sequence 674 BP; 128 A; 240 C; 208 G; 98 T; 0 other;
 Query Match 70.6%; Score 536.2; DB 21; Length 674;
 Best Local Similarity 98.2%; Pred. No. 5.1e-95;
 Matches 583; Conservative 0; Mismatches 78; Indels 0; Gaps 0;
 QY 18 CCGTCTCTGTGCTGTGACATGACGCTCCACCTGCTGTGTCTGTCTGTCTGTGT 77
 Db 14 CCGTCTCTCTGTGACATGACGCTCTCTGACGCTCTCTGACGCTCTCTGCTGTGTGT 73
 QY 78 ACAACAGGCTTTCGT 137
 Db 74 ACATGACGCTTTCGT 133
 QY 138 GGAATCATTCCTCCGAGCTCGGAGTACCCGAGCTTCACCGAGCTGGAGAACAA 197
 Db 134 AGAAATCATTCCTCCGAGCTCGGAGTACCCGAGCTTCCTGAGAACTTAGAACAACA 193
 QY 198 GACCATGAACCGGCGGAGAAACGAGGCGGCTCCGACACCCCTTTGAGACCAAGA 257
 Db 194 GACCATGAACCGGCGGAGAAACGAGGCGGACCCGACCATCTTATGACCAAGA 253
 QY 258 CGTGTCCGATGACGCTGCGGAGCTGACCTTACCCGCTTACGTAACGATGGCCGTG 317
 Db 254 CGTGTCCGATGACGCTGCGGAGCTGACCTTACCCGCTTACGTAACGATGGCCGTG 313
 QY 318 CCGGAGCGGCAAGCGGACCGGACCGAGCTGTGTGTCTCGGCGGAGCGGCGGCT 377
 Db 314 CCGGAGCGGCAAGCGGACCGGACCGAGCTGTGTGTCTCGGCGGAGCGGCGGCT 373
 QY 378 GCTGCCAAGCGCATCGGCGGCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 437
 Db 374 GCTGCCAAGCGCATCGGCGGCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 433
 QY 438 CATCCCGGAGCGGCTACCGGCGGAGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 497
 Db 434 CATCCCGGAGCGGCTACCGGCGGAGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 493
 QY 498 GCGGCGGCAAGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 557
 Db 494 GCGGCGGCAAGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 553
 QY 558 CAACGATCGAGCTCAAGGACTTTCGAGACCGAGCGCTGCGGCGGAGGAGGCGGAA 617
 Db 554 CAACGATCGAGCTCAAGGACTTTCGAGACCGAGCGCTGCGGCGGAGGAGGAGGAA 613
 QY 618 GCGGCGGCGGCGGCGGCGGAGCGGCAAGGCAACGAGCGGAGCTTGAGAAACGCTTACTA 677
 Db 614 GCGGCGGCGGCGGCGGCGGAGCGGCAAGGCAACGAGCGGAGCTTGAGAAACGCTTACTA 673
 QY 678 G 678
 Db 674 G 674
 RESULT 12
 ID AAA94049 standard; DNA; 5680 BP.
 XX AAA94049;
 AC AAA94049;
 DT 15-JAN-2001 (first entry)
 XX
 DE Human DAN/Cerberus-related protein 6 (hDCR6) gene.
 XX
 KM Human; DNA/Cerberus-related protein 6; hDCR6; morphogenic protein;

PT	/note= "This polymorphism results in an amino acid
FT	change from Arg to Xaa (stop codon)"
FT	replace (2190, T)
FT	/*tag= j
FT	/note= "This polymorphism results in an amino acid
FT	change from Trp to Xaa (stop codon)"
FT	replace (2408, G)
FT	/*tag= k
FT	replace (3539, T)
FT	/*tag= l
FT	replace (3944, T)
FT	/*tag= m
FT	replace (4425, G)
FT	/*tag= n
FT	replace (5097, A)
FT	/*tag= o
FT	complement (5099..2342)
FT	/*tag= p
FT	complement (5100..5331)
FT	/*tag= q
FT	/number= 1
FT	replace (5292, T)
FT	/*tag= r
FT	complement (5366..5405)
FT	/*tag= s
FT	replace (5666, A)
FT	/*tag= t
FT	replace (5696, T)
FT	/*tag= u
FT	replace (5833, C)
FT	/*tag= v
FT	replace (6070, G)
FT	/*tag= w
FT	replace (6392, G)
FT	/*tag= x
FT	replace (6474, G)
FT	/*tag= y
FT	replace (6718..6719, CTCCT)
FT	/*tag= z
FT	replace (6737, A)
FT	/*tag= aa
FT	replace (6867, C)
FT	/*tag= ab
FT	replace (6926, G)
FT	/*tag= ac
FT	replace (6927, T)
FT	/*tag= ad
XX	
XX	WO200198491-A2.
XX	
XX	27-DEC-2001.
XX	
XX	15-JUN-2001; 2001WO-EP06795.
XX	
XX	19-JUN-2000; 2000EP-0112867.
XX	
XX	(HOFF) HOFFMANN LA ROCHE & CO AG F.
XX	(UYIN-) UNIV INSTELLING ANTWERPEN UIA.
XX	
XX	Balemans W, Ebeling M, Foernzler D, Patel N, Van Hul W,
XX	Vickery BH;
XX	
XX	WPI; 2002-139789/18.
XX	
XX	P-PSDB; AAEL7089.
XX	
XX	Novel genetic polymorphisms in the Van Buchem-sclerosteosis disease
XX	region that are associated with abnormal bone formation useful for
XX	diagnosis and assessment of osteoporosis or sclerosteosis in humans
XX	
XX	Claim 11; Fig 3; 70pp; English.
XX	
XX	The invention relates to a nucleic acid encoding osteolevin region
XX	polymorphisms. The invention also relates to genetic polymorphisms in
CC	

CC the Van Buchem-sclerosteosis disease region that are associated with
CC disorders resulting in either net excess bone formation or insufficient
CC bone formation in humans. Osteolevin DNA is useful for screening for
CC osteolevin polymorphisms associated with abnormal bone formation in a
CC subject and for the presence of a heritably linked form of abnormal bone
CC formation in a subject, by determining the presence of a polymorphism in
CC the osteolevin nucleic acid sequence obtained from the subject.
CC Osteolevin protein is useful for treating diseases associated with
CC abnormal bone formation, such as sclerosteosis, Van Buchem's disease and
CC Paget's disease. Nucleic acids which encode genes in the osteolevin
CC region or their modified forms can also be used to generate either
CC transgenic animals or knockout animals which are useful in the screening
CC and development of therapeutically useful reagents. Osteolevin proteins
CC are useful in pharmacological characterisation of novel modulators of the
CC activity of protein and protein complexes. Human osteolevin gene is
CC located on chromosome 17. The present sequence is human osteolevin gene.
XX
SQ Sequence 7089 BP; 1530 A; 1944 C; 1928 G; 1697 T; 0 other;

Query Match	68.1%	Score 516.8;	DB 24;	Length 7099;
Best Local Similarity	96.0%	Pred. No. 2.9e-91;		
Matches 530; Conservative	0;	Mismatches 22;	Indels 0;	Gaps 0;

QY	208	CGGCGGAGAAACGAGAGGAGGAGCTTCCCCACACACCCCTTTAGAACCAAGAAGTGTCCGAG	267
Db	2390	CAGGGGTGGCCAGGCGGAGCGGACACCTTACAGCGGCGCTCTCTCCACAGAGTGTCCGAG	2331
QY	268	TACAGCTGACGCGAGCTGCACCTTACCCGCTTACGTGACCGATAGGGGCGTGTCCGAGGCGC	327
Db	2330	TACAGCTGTCCGCGAGCTGTGCACCTTACCCGCTTACGTGACCGATAGGGGCGTGTCCGAGGCGC	2271
QY	328	AAGCGGCTACCGAGCTGTGTGCTCCGAGCAGTGCAGCCCGGCGGCTGTCTGTCCCAAC	387
Db	2270	AAGCGGCTACCGAGCTGTGTGCTCCGAGCAGTGCAGGCGGCGGCGCTGTCTGTCCCAAC	2211
QY	388	GCCATCGGCGCGGCGCAAGTGTGTGGCCACTAGTGGGCGCGACTTCCGTGTGATCCCGAC	447
Db	2210	GCCATCGGCGCGGCGCAAGTGTGTGGCCACTAGTGGGCGCGACTTCCGTGTGATCCCGAC	2151
QY	448	CGCTACCGCGCGGCGAGCGCGTGCAGCTGTGTGCCCGGAGTGAAGGCGCGCGCGCGCG	507
Db	2150	CGCTACCGCGCGGCGAGCGCGTGCAGCTGTGTGCCCGGAGTGAAGGCGCGCGCGCGCGCG	2091
QY	508	AAGGTGCGGCTGTGTGAGCTTGTGTCAGAGTCAGAGCGGCTTACCCGCTTCCACACCAAGTCG	567
Db	2090	AAGGTGCGGCTGTGTGAGCTTGTGTCAGAGTCAGAGCGGCTTACCCGCTTCCACACCAAGTCG	2031
QY	568	GAGCTCAAGGACTTTCGGGACCGAGGCGCGCTCGGCGCGAGAGGGGCGGAGAGCGCGCGCC	627
Db	2030	GAGCTCAAGGACTTTCGGGACCGAGGCGCGCTCGGCGCGAGAGGGGCGGAGAGCGCGCGCC	1971
QY	628	CGGCGCGGAGGCGGCAAAAGCCACACAGGCGCGAGCTGTGAGGAAGCGCTTACGAGGCGCGCC	687
Db	1970	CGGCGCGGAGGCGGCAAAAGCCACACAGGCGCGAGCTGTGAGGAAGCGCTTACGAGGCGCGCC	1911
QY	688	GCGCGCTTCCCAACCGGCGGCGCGCCCGGCGCTTGAACCCGCGCGCCACATTTCTGTCTC	747
Db	1910	GCGCGCTTCCCAACCGGCGGCGCGCCCGGCGCTTGAACCCGCGCGCCACATTTCTGTCTC	1851
QY	748	TGCGGCTGTGTTT 759	
Db	1850	TGCGGCTGTGTTT 1839	

RESULT 14
AAA29064
ID AAA29064 standard; DNA; 9301 BP.
XX
AC AAA29064;
XX
DT 12-SEP-2000 (first entry)
DE Human TGF-beta binding protein (BEER) genomic DNA

XX osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
 KW BEER; chromosome 17q12-21; gene therapy; antisense therapy; fracture;
 XX bone mineralization; ds.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT exon 162..428
 FT /*tag= a
 FT /number= 1
 FT CDS 209..3608
 FT /*tag= b
 FT /product= TGF-beta_binding_protein
 FT intron 429..3185
 FT /*tag= c
 FT exon 3186..5219
 FT /*tag= d
 FT /number= 2
 XX
 PN WO200032773-A1.
 XX
 PD 08-JUN-2000.
 XX
 PF 24-NOV-1999; 99WO-US27990.
 XX
 PR 27-NOV-1998; 98US-0110283.
 XX
 PA (DARW-) DARWIN DISCOVERY LTD.
 XX
 PI Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepfer BW,
 PI Van Ness J, Winkler DG;
 DR WPI; 2000-412321/35.
 XX
 DR P-PSDB; AAY96429.
 XX
 PT Nucleic acids (1) encoding a transforming growth factor beta binding
 PT protein, useful for identifying agents for treating osteopenia,
 PT osteoporosis and fractures
 XX
 PS Disclosure; Page 87-96; 162pp; English.
 XX
 CC This DNA encodes a human transforming growth factor-beta (TGF-beta)
 CC binding protein designated BEER. The hBEER gene has been localized
 CC to the chromosome 17q12-21. The cDNA and protein may be used for
 CC prevention, treatment and diagnosis of diseases associated with
 CC inappropriate BEER expression. For example, they may be used to treat
 CC disorders associated with decreased TGF-beta BP expression. The cDNA or
 CC vectors may be administered to treat diseases by rectifying mutations or
 CC deletions in a patient's genome that affect the activity of BEER by
 CC expressing inactive proteins or to supplement the patients own production
 CC of BEER polypeptides. The nucleic acids may be used for recombinant
 CC production of BEER, gene therapy, antisense therapy, as probes for
 CC diagnostic assays and for functional studies. BEER may be used to raise
 CC antibodies and for identification of BEER modulators. BEER antagonists
 CC may be used to increase bone mineral content for the treatment of
 CC disorders such as osteopenia, osteoporosis, fractures and other
 CC disorders associated with low mineral content.
 XX
 SQ Sequence 9301 BP; 2276 A; 2479 C; 2503 G; 2043 T; 0 other;
 XX
 Query Match 68.1%; Score 516.8; DB 21; Length 9301;
 Best Local Similarity 96.0%; Pred. No. 2.9e-91;
 Matches 530; Conservative 0; Mismatches 22; Indels 0; Gaps 0;
 XX
 QY 208 CCGGGCGGAGAAAGGAGGGGGCTCCACACACCCCTTTGAACCAAGAGTGTCCGAG 267
 DB 3138 CAGGGGTGGCCAGGGCGGGCCACCCCTCCTCTCCACAGACGTGTCCGAG 3197
 QY 268 TACAGTGGCGGAGCTGCATTCACCCGCTACGTACCGATGGGCGGCGGAGCGCC 327
 DB 3198 TACAGTGGCGGAGCTGCATTCACCCGCTACGTACCGATGGGCGGCGGAGCGCC 3257

QY 328 AAGCGGTACACGAGCTGTGTGCTCCGGCCAGTGGCGCCCGGCGCTGTGCCCAAC 387
 DB 3258 AAGCGGTACACGAGCTGTGTGCTCCGGCCAGTGGCGCCCGGCGCTGTGCCCAAC 3317
 QY 388 GCCATGGCCCGGCGCAAGTGTGGCGACCTAGTGGGCCCGACCTTCCGCTGATCCCGGAC 447
 DB 3318 GCCATGGCCCGGCGCAAGTGTGGCGACCTAGTGGGCCCGACCTTCCGCTGATCCCGGAC 3317
 QY 448 CGCTACCGCGGCGAGCGGCTGAGCTGTGTCCCGGTGTGAGGCGCGCGCGCGC 507
 DB 3378 CGCTACCGCGGCGAGCGGCTGAGCTGTGTCCCGGTGTGAGGCGCGCGCGCGC 3437
 QY 508 AAGTGGCGCTGTGTGTGCTGTGCAAGTGAAGCGCTTCAACCAACAGTGG 567
 DB 3438 AAGTGGCGCTGTGTGTGCTGTGCAAGTGAAGCGCTTCAACCAACAGTGG 3497
 QY 568 GAGCTCAAGGACTTTCGGGACCGGAGCGCTTGGCGCGCAAGAGGCGCGAGCCCGGCCC 627
 DB 3498 GAGCTCAAGGACTTTCGGGACCGGAGCGCTTGGCGCGCAAGAGGCGCGAGCCCGGCCC 3557
 QY 628 CGCGCCCGGAGCGGCAAGCGCAACGAGCGGAGCTGAGAGAGCGCTTCTAGAGCGCGCC 687
 DB 3558 CGCGCCCGGAGCGGCAAGCGCAACGAGCGGAGCTGAGAGAGCGCTTCTAGAGCGCGCC 3617
 QY 688 GCGCCCTCTCCACCGCGGCGCGCCCGGCGCTTGAACCGCGCCCAATTCTGTCTTC 747
 DB 3618 GCGCCCTCTCCACCGCGGCGCGCCCGGCGCTTGAACCGCGCCCAATTCTGTCTTC 3677
 QY 748 TCGCGGTGTGTTT 759
 DB 3678 TCGCGGTGTGTTT 3689

RESULT 15
 ID AAA29058
 ID AAA29058 standard; cDNA; 638 BP.
 AC AAA29058;
 XX
 DT 12-SEP-2000 (first entry)
 XX
 DE Murine TGF-beta binding protein (BEER) cDNA.
 XX
 KW osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
 KW BEER; gene therapy; antisense therapy; fracture; bone mineralization; se.
 XX
 OS Mus musculus.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..636
 FT /*tag= a
 FT /product= TGF-beta_binding_protein
 XX
 PN WO200032773-A1.
 XX
 PD 08-JUN-2000.
 XX
 PF 24-NOV-1999; 99WO-US27990.
 XX
 PR 27-NOV-1998; 98US-0110283.
 XX
 PA (DARW-) DARWIN DISCOVERY LTD.
 PI Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepfer BW,
 PI Van Ness J, Winkler DG;
 DR WPI; 2000-412321/35.
 DR P-PSDB; AAY96432.
 XX
 PT Nucleic acids (1) encoding a transforming growth factor beta binding
 PT protein, useful for identifying agents for treating osteopenia,
 PT osteoporosis and fractures
 XX

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 28, 2003, 18:49:00 ; Search time 90.8624 Seconds
(without alignments)
7108.427 Million cell updates/sec

Title: US-09-867-274-1

Perfect score: 759
Sequence: 1 taccgaagcgcgcgtccccc.....ctgcctcgcgcgtgttt 759

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 574371 seqs, 425486471 residues

Total number of hits satisfying chosen parameters: 1148742

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA.*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCT_US_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	759	100.0	759	US-09-867-274-1	Sequence 1, Appli
2	501.6	66.1	636	US-09-867-274-3	Sequence 3, Appli
3	422	55.6	422	US-09-864-761-30988	Sequence 30988, A
4	355	46.8	392	US-09-864-761-14440	Sequence 14440, A
5	82.6	10.9	954	US-10-152-661-430	Sequence 430, App
6	82.6	10.9	954	US-09-866-050A-430	Sequence 430, App
7	82.6	10.9	962	US-10-152-661-40	Sequence 40, Appl
8	82.6	10.9	962	US-10-152-661-209	Sequence 209, App
9	82.6	10.9	962	US-09-866-050A-40	Sequence 40, Appl
10	82.6	10.9	962	US-09-866-050A-209	Sequence 209, App
11	69.6	9.2	900	US-09-853-625B-1	Sequence 1, Appli
12	69.6	9.2	1692	US-09-745-288-92	Sequence 92, Appli
13	52	6.9	4020	US-09-796-679-5	Sequence 5, Appli
14	49.6	6.5	4257	US-09-825-288A-1	Sequence 1, Appli
15	49	6.5	4098	US-09-962-436-37	Sequence 37, Appl
16	48.6	6.4	174424	US-09-967-768A-314	Sequence 314, App
17	47.4	6.2	2307	US-09-893-519A-87	Sequence 87, Appl
18	47	6.2	1614	US-09-976-740-45	Sequence 45, Appl
19	47	6.2	1614	US-10-023-529-45	Sequence 45, Appl

20	47	6.2	1614	US-10-023-523-45	Sequence 45, Appl
21	47	6.2	12425	US-09-976-740-50	Sequence 50, Appl
22	47	6.2	12425	US-10-023-529-50	Sequence 50, Appl
23	47	6.2	12425	US-10-023-523-50	Sequence 50, Appl
24	46.8	6.2	1910	US-09-789-836-1	Sequence 1, Appli
25	46.8	6.2	2198	US-10-044-205A-1	Sequence 1, Appli
26	46.8	6.2	2249	US-10-217-745-5	Sequence 5, Appli
27	46.8	6.2	3651	US-09-964-469-3	Sequence 3, Appli
28	44	5.8	14800	US-09-954-456-1601	Sequence 1601, Ap
29	44	5.8	88421	US-09-976-059-1	Sequence 1, Appli
30	43.8	5.8	1275	US-09-815-242-4042	Sequence 4042, Ap
31	43.6	5.7	440	US-10-184-644-202	Sequence 202, App
32	43.6	5.7	8459	US-09-817-913-8	Sequence 8, Appli
33	43.6	5.7	8459	US-09-817-538-8	Sequence 8, Appli
34	43.4	5.7	8867	US-09-932-367A-22	Sequence 22, Appl
35	43.4	5.7	88421	US-09-976-059-1	Sequence 1, Appli
36	43.2	5.7	13023	US-09-815-242-4052	Sequence 4052, Ap
37	42.8	5.7	33023	US-09-880-107-3350	Sequence 30, Appl
38	42.8	5.6	13842	US-09-860-846-30	Sequence 30, Appl
39	42.8	5.6	13842	US-09-861-289-30	Sequence 30, Appl
40	42.8	5.6	36778	US-09-860-846-5	Sequence 5, Appli
41	42.8	5.6	36778	US-09-861-289-5	Sequence 5, Appli
42	42.2	5.6	5173	US-09-880-107-3356	Sequence 3356, Ap
43	42.2	5.6	12606	US-09-957-974-2	Sequence 2, Appli
44	42.2	5.6	17862	US-10-092-154-1313	Sequence 1313, Ap
45	42.2	5.6	17862	US-09-764-847-1313	Sequence 1313, Ap

ALIGNMENTS

RESULT 1
US-09-867-274-1
Sequence 1, Application US/09867274
Patent No. US20020106650A1
GENERAL INFORMATION:
APPLICANT: Paszty, Christopher
TITLE OF INVENTION: Cysteine Knot Polypeptides: Cloaked-2 Molecules and Uses Thereof
FILE REFERENCE: 0101737428
CURRENT APPLICATION NUMBER: US/09/867,274
CURRENT FILING DATE: 2001-05-29
PRIOR APPLICATION NUMBER: US 60/208,550
PRIOR FILING DATE: 2000-06-01
PRIOR APPLICATION NUMBER: US 60/223,542
NUMBER OF SEQ ID NOS: 25
SOFTWARE: PatentIn version 3.0
SEQ ID NO 1
LENGTH: 759
TYPE: DNA
ORGANISM: Homo sapiens
US-09-867-274-1

Query Match 100.0%; Score 759; DB 10; Length 759;
Best Local Similarity 100.0%; Pred. No. 1.2e-184;
Matches 759; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TACGGAAGTGGGCTCCCTCTGCTGTAACATGACGCTCCCATGCGCTCTGT 60
1 TACTGGAAGTGGGCTCCCTCTGCTGTAACATGACGCTCCCATGCGCTCTGT 60
DB 1 TACTGGAAGTGGGCTCCCTCTGCTGTAACATGACGCTCCCATGCGCTCTGT 60
QY 61 CTCGCTCCTCGTGTGACACAGCCTTCGTAAGTGAAGGCGCAGGGGTGACAGG 120
61 CTCGCTCCTCGTGTGACACAGCCTTCGTAAGTGAAGGCGCAGGGGTGACAGG 120
DB 61 CTCGCTCCTCGTGTGACACAGCCTTCGTAAGTGAAGGCGCAGGGGTGACAGG 120
QY 121 TTCAAGATGATGCCAGGAATCATCCCGAGCTCGAGAGTACCCCGAGCTTCAACG 180
121 TTCAAGATGATGCCAGGAATCATCCCGAGCTCGAGAGTACCCCGAGCTTCAACG 180
DB 121 TTCAAGATGATGCCAGGAATCATCCCGAGCTCGAGAGTACCCCGAGCTTCAACG 180
QY 181 GAGCTGGAACAACAAGACCATGAACCGGCGGAGAAAGGCGCGCTCCCAACAC 240
181 GAGCTGGAACAACAAGACCATGAACCGGCGGAGAAAGGCGCGCTCCCAACAC 240
DB 181 GAGCTGGAACAACAAGACCATGAACCGGCGGAGAAAGGCGCGCTCCCAACAC 240

QY 241 CCCTTTGAGACCAAGAGCTGTCCGAGTACAGCTCCGCGAGCTGACCTTACACCCCGTAC 300
|
|
|
Db 241 CCCCTTTAGAACAAAGAGCTGTCCGAGTACAGCTCCGCGAGCTGACCTTACACCCCGTAC 300
|
|
|
QY 301 GTGACCGATGGGCGGTGCGCGAGCGCGCAAGCGGTGTCAACGAGTGTGTCTCCGCGCAG 360
|
|
|
Db 301 GTGACCGATGGGCGGTGCGCGAGCGCGCAAGCGGTGTCAACGAGTGTGTCTCCGCGCAG 360
|
|
|
QY 361 TGGCGCCCGGCGCGCTGTGCGCAACGCGCATGCGCGCGCAAGTGTGTGCGCACTTGT 420
|
|
|
Db 361 TGGCGCCCGGCGCGCTGTGCGCAACGCGCATGCGCGCGCAAGTGTGTGCGCACTTGT 420
|
|
|
QY 421 GGGGCGGACCTTCCGCTGTGATCCCGACCGCGACCGCGCGAGCGCGGTGCGAGCTGTGT 480
|
|
|
Db 421 GGGGCGGACCTTCCGCTGTGATCCCGACCGCGACCGCGCGAGCGCGGTGCGAGCTGTGT 480
|
|
|
QY 481 CCCGATGATGAGCGCGCGCGCGCAAGGATGCGCTGTGTGCGCTGTGTCAAGTGTCAAG 540
|
|
|
Db 481 CCCGATGATGAGCGCGCGCGCGCAAGGATGCGCTGTGTGCGCTGTGTCAAGTGTCAAG 540
|
|
|
QY 541 CGCCTCACCGGCTTCCACACCAAGTGTGAGTCAAGGATTTGCGGACCGAGCGCGCTCGG 600
|
|
|
Db 541 CGCCTCACCGGCTTCCACACCAAGTGTGAGTCAAGGATTTGCGGACCGAGCGCGCTCGG 600
|
|
|
QY 601 CCGGAGAAAGGCGCGGAGCGCGCGCGCGCGCGCGGCGCGGAGCGCAACCGAGCGCGAG 660
|
|
|
Db 601 CCGGAGAAAGGCGCGGAGCGCGCGCGCGCGCGCGGCGCGGAGCGCAACCGAGCGCGAG 660
|
|
|
QY 661 CTGAGAAAGCGCTTCTAGAGCGCGCGCGCGCGCTTCCCAACCGGCGCGCGCGCGCT 720
|
|
|
Db 661 CTGAGAAAGCGCTTCTAGAGCGCGCGCGCGCGCTTCCCAACCGGCGCGCGCGCGCT 720
|
|
|
QY 721 GAACCCCGCGCGCGCAATTTCTGTCTCTGTGCGGTGTGT 759
|
|
|
Db 721 GAACCCCGCGCGCGCAATTTCTGTCTCTGTGCGGTGTGT 759
|
|
|

RESULT 2

US-09-867-274-3
; Sequence 3, Application US/09867274
; Patent No. US20020106650A1
; GENERAL INFORMATION:
; APPLICANT: Paetzly, Christopher
; APPLICANT: Geo, Xongming
; TITLE OF INVENTION: Cysteine Knot Polypeptides: Cloaked-2 Molecules and Uses Thereof
; FILE REFERENCE: 01017/37428
; CURRENT APPLICATION NUMBER: US/09/867,274
; PRIOR FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: US 60/208,550
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: US 60/223,542
; PRIOR FILING DATE: 2000-08-04
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 636
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-867-274-3

Query March 66.1%; Score 501.6; DB 10; Length 636;
Best Local Similarity 87.5%; Pred. No. 4.3e-119;
Matches 562; Conservative 0; Mismatches 74; Indels 6; Gaps 1;

QY 37 ATGAGAGTCCCACTGCGCCTGTGTCTGCTGTGCTGTGATACACAGAGCTTCCGTGTA 96
|
|
|
Db 1 ATGAGAGTCCCACTGCGCCTGTGTCTGCTGTGCTGTGATACACAGAGCTTCCGTGTA 96
|
|
|
QY 97 GTGAGAGGCGAGGGGTGCGAGCGCTTCAAGATGATGCGCAAGATATCATCCCGAGCTC 156
|
|
|
Db 61 GTGAGAGGCGAGGGGTGCGAGCGCTTCAAGATGATGCGCAAGATATCATCCCGAGCTC 120
|
|
|

QY 157 GGAGAGTACCCCGAGCTCTCCACCGAGCTGGAGAACAAAGACCATGAACCGGCGGAG 216
|
|
|
Db 121 GGAGAGTACCCCGAGCTCTCTCC-----TGAGAACAAAGACCATGAACCGGCGGAG 174
|
|
|
QY 217 AACGAGGAGCGGCGCTCCCGACACCGCTTTGAGACCAAGAGCTGTCCAGTACAGCTGC 276
|
|
|
Db 175 AATGAGGAGAGACTCCCGACCATCTATGACCGCAAGAGTGTGTGCGAGTACAGCTGC 234
|
|
|
QY 277 CGCGAGCTGACACTTACCCCGTCTAGTGAACGATGGGCGGTGCGCGAGCGCAACCGGTC 336
|
|
|
Db 235 CGCGAGCTGACACTTACCCCGTCTAGTGAACGATGGGCGGTGCGCGAGCGCAACCGGTC 294
|
|
|
QY 337 ACCGAGCTGTGTGCTCCGCGCAGTGTGCGCGCGCGCGCGCTGTGCGCAACCGCATGCGC 396
|
|
|
Db 295 ACCGAGTGTGTGCTCCGCGCAGTGTGCGCGCGCGCGCGCTGTGCGCAACCGCATGCGC 354
|
|
|
QY 397 CGCGCAAGTGTGCGCACTTGTGTGCGCGCGCACTTCCGCTGTGATCCCGACCGCTACCGC 456
|
|
|
Db 355 CGCGTGAAGTGTGCGCGCGCGCGCAACCGGATTTCCGCTGTGATCCCGATCGCTACCGC 414
|
|
|
QY 457 GGGCAGCGCGGTGACGCTGTGTCCGCTGTGAGAGCGCGCGCGCGCAAGTGTGCGC 516
|
|
|
Db 415 GGGCAGCGCGGTGACGCTGTGTCCCGGCGCGCGCGCGCGCGCTGTGCGCAAGTGTGCGC 474
|
|
|
QY 517 CTGTGTGCTCTGTGCAAGTGTGAGCGCTTCAACCGCTTCCCAACCAAGTGTGAGCTCAAG 576
|
|
|
Db 475 CTGTGTGCTCTGTGCAAGTGTGAGCGCTTCAACCGCTTCCCAACCAAGTGTGAGCTCAAG 534
|
|
|
QY 577 GACTTGTGGGACCGAGCGCGCTGTGCGCGCGCAAGAGGCGCGGCGCGCGCGCGCGCG 636
|
|
|
Db 535 GACTTGTGGGACCGAGCGCGCGCGCGCGCGCGCAAGAGGCTGTGCGCAAGCGCGCGCGCGCG 594
|
|
|
QY 637 AGCGCAAGCAACCAAGCGGAGCTGTGAGAGCGCTTCAAG 678
|
|
|
Db 595 GAGCGCAAGCAACCAAGCGGAGCTGTGAGAGCGCTTCAAG 636
|
|
|

RESULT 3

US-09-864-761-30988/C
; Sequence 30988, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wenheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecmika-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30


```

; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 30988
; LENGTH: 422
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC003098.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.62
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.6
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.49
; OTHER INFORMATION: EST HUMAN HIT: BE613498.1, EVALUE 9.90e-02
; OTHER INFORMATION: SWISSPROT HIT: P45646, EVALUE 4.90e-01
; OTHER INFORMATION: NT HIT: AF074705.1, EVALUE 1.00e+00
; US-09-864-761-30988

```

```

Query Match      55.6%; Score 422; DB 10; Length 422;
Best Local Similarity 100.0%; Pred. No. 8.3e-99;
Matches 422; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 257 ACCTGTCAGTACAGTCCGCGGAGTGTGACTTCAACCCGCTACGAGCCGATGAGCCCT 316
DB 422 ACCTGTCAGTACAGTCCGCGGAGTGTGACTTCAACCCGCTACGAGCCGATGAGCCCT 363
QY 317 GCCGAGCGCCAGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 376
DB 362 GCCGAGCGCCAGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 303
QY 377 TGTGTCCTCAACGCGCATGCGCGGCGGCAAGTGTGCGGAGCTTACGATGAGCCCT 436
DB 302 TGTGTCCTCAACGCGCATGCGCGGCGGCAAGTGTGCGGAGCTTACGATGAGCCCT 243
QY 437 GCATCCCCCAACCCGCTACCGCGGCGGCAAGTGTGCGGAGCTTACGATGAGCCCT 496
DB 242 GCATCCCCCAACCCGCTACCGCGGCGGCAAGTGTGCGGAGCTTACGATGAGCCCT 183
QY 497 CGCGCGCGCGCAAGTGTGCGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGG 556
DB 182 CGCGCGCGCGCAAGTGTGCGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGG 123
QY 557 ACAACAGTGTGAGTCAAGAGCTTGTGAGGCGCGGCTGTGAGGCGCGCAAGAGGCGCGGA 616
DB 122 ACAACAGTGTGAGTCAAGAGCTTGTGAGGCGCGGCTGTGAGGCGCGCAAGAGGCGCGGA 63
QY 617 AGCGCGCGCGCGCGCGCGCGCGCAAGCGCAAGCGCGCGCGCGCGCGCGCGCGCGCGCG 676
DB 62 AGCGCGCGCGCGCGCGCGCGCGCAAGCGCAAGCGCGCGCGCGCGCGCGCGCGCGCGCG 3
QY 677 AG 678
DB 2 AG 1

```

```

RESULT 4
US-09-864-761-14440/c
; Sequence 14440, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.

```

```

; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aemica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 14440
; LENGTH: 392
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC003098.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.62
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.6
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.49
; US-09-864-761-14440

```

```

Query Match      46.8%; Score 355; DB 10; Length 392;
Best Local Similarity 100.0%; Pred. No. 9.6e-82;
Matches 355; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 405 GTGTGTGAGTGTGAGGCGCGGATTCGCTGTGATCCCGGACGCTTACCGCGGCAAGCG 464
DB 392 GTGTGTGAGTGTGAGGCGCGGATTCGCTGTGATCCCGGACGCTTACCGCGGCAAGCG 333
QY 465 CGTGTGAGTGTGTCTCCCGGTGTGAGGCGCGGCGCGGCAAGTGTGAGGCGCTGTGTGAGG 524
DB 332 CGTGTGAGTGTGTCTCCCGGTGTGAGGCGCGGCGCGGCAAGTGTGAGGCGCTGTGTGAGG 273
QY 525 CTGTGTGAGTGTGAGGCGCGGCTTCCACCAACGAGTGTGAGGCTTCAAGGAGCTTGG 584
DB 272 CTGTGTGAGTGTGAGGCGCGGCTTCCACCAACGAGTGTGAGGCTTCAAGGAGCTTGG 213

```

QY 585 GACGAGGCGCTGCGCGCGAGAGGCGGAGCCGCGCGCGCGAGAGCCCA 644
| | | | |
DB 212 GACGAGGCGCTGCGCGCGAGAGGCGGAGCCGCGCGCGCGAGAGCCCA 133
| | | | |
QY 645 AGCAACGAGCGCGAGCTGAGAAAGCCTACTAGAGCCGCGCGCGCGCTCCACCG 704
| | | | |
DB 152 AGCAACGAGCGCGAGCTGAGAAAGCCTACTAGAGCCGCGCGCGCGCTCCACCG 93
| | | | |
QY 705 CGGCGCGCGCGCGCGCTGAGAAAGCCTACTAGAGCCGCGCGCGCGCTCCACCG 759
| | | | |
DB 92 CGGCGCGCGCGCGCGCTGAGAAAGCCTACTAGAGCCGCGCGCGCGCTCCACCG 38
| | | | |

RESULT 5

US-10-152-661-430
; Sequence 430, Application US/10152661
; Publication No. US2003022835A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; TITLE OF INVENTION: and Methods for Their Use
; FILE REFERENCE: 11000.1011c5
; CURRENT APPLICATION NUMBER: US/10/152.661
; CURRENT FILING DATE: 2002-05-20
; PRIOR APPLICATION NUMBER: 09/866,050
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 60/221,232
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: 60/206,650
; PRIOR FILING DATE: 2000-05-24
; PRIOR APPLICATION NUMBER: 09/312,283
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: PCT/NZ99/00051
; PRIOR FILING DATE: 1999-04-29
; PRIOR APPLICATION NUMBER: 09/188,930
; PRIOR FILING DATE: 1998-11-09
; PRIOR APPLICATION NUMBER: 09/069,726
; PRIOR FILING DATE: 1998-04-29
; NUMBER OF SEQ ID NOS: 725
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 430
; LENGTH: 954
; TYPE: DNA
; ORGANISM: Mouse
US-10-152-661-430

Query Match 10.9%; Score 82.6; DB 9; Length 954;
Best Local Similarity 52.5%; Pred. No. 2.2e-12;
Matches 287; Conservative 0; Mismatches 239; Indels 21; Gaps 4;
QY 112 TGGCAGGCGCTTCAAGATGATGTCACGGAATCATCCCGAGCTCGAGAGTACCCCGAG 171
| | | | |
DB 99 TGTGTGGCTTTTAAATGATGTCACGGAATCTTTATTACATGTGTAACTGTCTC 158
| | | | |
QY 172 CTTCCACGAGCTGAGAACAAAGCATGAACCGGCGGAGAAAGGAGGCGGCT 231
| | | | |
DB 159 CCGGAGC---ACCCAGAGCAAGCAGCCTGATCAAGCAAGAAATGAGGAGGAGCAT 215
| | | | |
QY 232 CCCACACCC---CTTGAAGCAAAAGAGTGTCCAGTACAGTCCCGGAGCTGCAC 288
| | | | |
DB 216 TTCAAGTACCTGAGCTGATGCAAAAGTCAAGTTCAGTGGGCTGAGGAACTCCGG 275
| | | | |
QY 289 TTCAACCGCTAGTACCGATGAGCGGCGGAGCGGCAAGCGGCTGACCGAGCTGTG 348
| | | | |
DB 276 TCCACCAATATATTTTGGAGCGGCGGAGGAGCATCAAGCTCTTGAAGAGAGCTGTG 335
| | | | |
QY 349 TGTCTCCGAGTGTGCGGCGGCGGCTGTGCTCCAAAGCCATCGGCGCGGCGG--- 402
| | | | |

DB 336 TGGCGGCGGAGTGTGCTTCCCTGCGGCTGCTTCCCACTGATGAGAGGCTACGGA 395
| | | | |
QY 403 -----AGTGTGTGCGACACTAGTGGGCGCGACTTCGCTGATCCCGACCGTACCGC 456
| | | | |
DB 396 ACAAGTACTGAGAGCGGAGAGGCTCTAGAGTGGCGGTGTGTCAAGCAAGAGCGGC 455
| | | | |
QY 457 GCGAGCGGCTGACGCTGTGTGTCCGATGTGAGGCGCGCGCGCGCGGAGGTGGC 516
| | | | |
DB 456 ACCAGAGGATGCTACGCTGAGTGTGAGAGCG---CAGCAGCGGACCTCAAAATACCC 512
| | | | |
QY 517 GTGTGTGCTGTGTGCAAGTGTGAGCGCTCAACCGGCTTCCAAACGATCGAGCTCAAG 576
| | | | |
DB 513 GTGTGTGCTGTGTGCAAGTGTGAGCGCTCAACCGGCTTCCAAACGATCGAGCTCAAG 572
| | | | |
QY 577 GACTTCGAGACGAGCGGCTGCGCGCGGAGAGGCGGAGCGGCGCGCGCGCGCG 636
| | | | |
DB 573 TTGAAAGCTGTGTGCGCAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 632
| | | | |
QY 637 AGCGCCA 643
| | | | |
DB 633 AATCCA 639
| | | | |

RESULT 6

US-09-866-050A-430
; Sequence 430, Application US/09866050A
; Publication No. US20030040471A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; TITLE OF INVENTION: and Methods for Their Use
; FILE REFERENCE: 11000.1011c4U
; CURRENT APPLICATION NUMBER: US/09/866.050A
; CURRENT FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 725
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 430
; LENGTH: 954
; TYPE: DNA
; ORGANISM: Mouse
US-09-866-050A-430

Query Match 10.9%; Score 82.6; DB 9; Length 954;
Best Local Similarity 52.5%; Pred. No. 2.2e-12;
Matches 287; Conservative 0; Mismatches 239; Indels 21; Gaps 4;
QY 112 TGGCAGGCGCTTCAAGATGATGTCACGGAATCATCCCGAGCTCGAGAGTACCCCGAG 171
| | | | |
DB 99 TGTGTGGCTTTTAAATGATGTCACGGAATCTTTATTACATGTGTAACTGTCTC 158
| | | | |
QY 172 CTTCCACGAGCTGAGAACAAAGCATGAACCGGCGGAGAAAGGAGGCGGCT 231
| | | | |
DB 159 CCGGAGC---ACCCAGAGCAAGCAGCCTGATCAAGCAAGAAATGAGGAGGAGCAT 215
| | | | |
QY 232 CCCACACCC---CTTGAAGCAAAAGAGTGTCCAGTACAGTCCCGGAGCTGTAC 288
| | | | |
DB 216 TTCAAGTACCTGAGCTGATGCAAAAGTCAAGTTCAGTGGGCTGAGGAACTCCGG 275
| | | | |
QY 289 TTCAACCGCTAGTACCGATGAGCGGCGGAGCGGCAAGCGGCTGACCGAGCTGTG 348
| | | | |
DB 276 TCCACCAATATATTTTGGAGCGGCGGAGGAGCATCAAGCTCTTGAAGAGAGCTGTG 335
| | | | |
QY 349 TGTCTCCGAGTGTGCGGCGGCGGCTGTGCTCCAAAGCCATCGGCGCGGCGG--- 402
| | | | |
DB 336 TGGCGGCGGAGTGTGCGGCGGCGGCTGTGCTCCAAAGCCATCGGAGAGGAGTACGGA 395
| | | | |
QY 403 -----AGTGTGTGCGACACTAGTGGGCGGAGCTTCCGCTGATCCCGACCGCTACCGC 456
| | | | |

Db 396 ACAAGTACTGAGCCGAGGAGAGCTCTCAGAGTGGCGGTGTCTCAACAACAAGCCGC 455
Qy 457 GCGCAGCGCGTGCAGCTGTGTCTCCGCGGTGTGAGAGCGCCGCGCGCAAGGTGGC 516
Db 456 ACCCAGAGATCCAGCTGTGTCTCAGAGCG--CAGCAGCGCGCACTTCAAAATCACC 512
Qy 517 CTGTGGCTGTGTGCAAGTGCAGAGCGCTTCACCCGCTTCACCAACAGTCCGAGCTCAAG 576
Db 513 GTGTGTCAGCGCGTGTGCAAGTGTACACCGTTCAGACAGAGTCCAGCCCAAC 572
Qy 577 GACTTGGAGCCAGCGCGCTCCGCGCGCAGAGGCGCGGCGCGCGCGCGCG 636
Db 573 TTGAAAGCGTGTGTGCGCAGCAAGCCCGCCAGCACCACAGAGAGCGGAAGAGCCAGC 632
Qy 637 AGCGCCA 643
Db 633 AATCCA 639

RESULT 7

US-10-152-661-40
; Sequence 40, Application US/10152661
; Publication No. US20030022835A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; FILE REFERENCE: 11000.1011CS
; CURRENT APPLICATION NUMBER: US/10/152,661
; CURRENT FILING DATE: 2002-05-20
; PRIOR APPLICATION NUMBER: 09/866,050
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 60/221,232
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: 60/206,650
; PRIOR FILING DATE: 2000-05-24
; PRIOR APPLICATION NUMBER: 09/312,283
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: PCT/NZ99/00051
; PRIOR FILING DATE: 1999-04-29
; PRIOR APPLICATION NUMBER: 09/188,930
; PRIOR FILING DATE: 1998-11-09
; PRIOR APPLICATION NUMBER: 09/069,726
; PRIOR FILING DATE: 1998-04-29
; NUMBER OF SEQ ID NOS: 725
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 40
; LENGTH: 962
; TYPE: DNA
; ORGANISM: Mouse
US-10-152-661-40

Query Match 10.9%; Score 82.6; DB 9; Length 962;
Best Local Similarity 52.5%; Pred. No. 2,2e-12;
Matches 287; Conservative 0; Mismatches 239; Indels 21; Gaps 4;

Qy 112 TGGCAGCGCTTCAAGATGATGCGAGGAATCATCCCGAGCTCGAGAGTACCCCGAG 171
Db 107 TGTGGCTTTTAAATATGATGCAAGAAATCTTTATTCATATGTGTAAACCTGTGC 166
Qy 172 CCTCCACCGAGTGTGAGAACAAAGACCATATACCGGCGGAGAAAGGAGCGCGCT 231
Db 167 CCGGAC--ACCCAGCAGCAACAGCAGCCCTGATCAAGCCGAGATGAGGAGCGCAT 223
Qy 232 CCCACACACCC--CTTGAAGACCAAGAGCTGTCCGAGTACAGCTGCGCGAGCTGCAC 288
Db 224 TTCAATGACACTGGAAGTGTGCAAAAGTTCAGATTCAGTGGGCTGACAGGAACTGGC 283

Qy 289 TTCACCCGCTACGTGACCGATAGGCGCGTGTCCGAGCGCCAAAGCCGCTGACCGAGCTGTG 348
Db 284 TTCACCAATATATTTCCGAGCGCGCAGAGTACACAGANTCAGCCCTCTTGAAGAGCTGGT 343
Qy 349 TGTCTCCGCTCAGTGTGCGCGCGCGCGCTGTGCTGCTCCAAAGCAGTATGCGCGCGC----- 402
Db 344 TGGCGGCGAGGTGCTTCCCTCCCTGCGGTGTCTTCCCACTGAATCGAGAGGAGTACGGA 403
Qy 403 -----AAGTGTGCGACCTAGTGGCGCCGACTTCCGCTGATCCCGACCGCTACCGC 456
Db 404 ACAAGTACTGAGCCGAGAGAGAGCTCTAGAGATGGCGGTGTGTCAACAACAAGCGCGC 463
Qy 457 GCGCAGCGCGTGCAGCTGTGTCTCCGCGGTGTGAGAGCGCGCGCGCGCGCAAGTGGC 516
Db 464 ACCCAGAGATCCAGCTGTGTCTCAGAGCG--CAGCAGCGCGCACTTCAAAATCACC 520
Qy 517 CTGTGGCTGTGTGCAAGTGCAGAGCGCTTCACCCGCTTCACCAACAGTCCGAGCTCAAG 576
Db 521 GTGTGTCAGCGCGTGTGCAAGTGTACACCGTTCAGACCAACAGATTCAGCCACAC 580
Qy 577 GACTTGGAGCCAGCGCGCTCCGCGCGCAGAAAGGCGCGCGCGCGCGCGCGCG 636
Db 581 TTGAAAGCGTGTGTGCGCAGCAAGCCCGCCAGCACCACAGAGAGCGGAAGAGCCAGC 640
Qy 637 AGCGCCA 643
Db 641 AATCCA 647

RESULT 8

US-10-152-661-209
; Sequence 209, Application US/10152661
; Publication No. US20030022835A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; FILE REFERENCE: 11000.1011CS
; CURRENT APPLICATION NUMBER: US/10/152,661
; CURRENT FILING DATE: 2002-05-20
; PRIOR APPLICATION NUMBER: 09/866,050
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 60/221,232
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: 60/206,650
; PRIOR FILING DATE: 2000-05-24
; PRIOR APPLICATION NUMBER: 09/312,283
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: PCT/NZ99/00051
; PRIOR FILING DATE: 1999-04-29
; PRIOR APPLICATION NUMBER: 09/188,930
; PRIOR FILING DATE: 1998-11-09
; PRIOR APPLICATION NUMBER: 09/069,726
; PRIOR FILING DATE: 1998-04-29
; NUMBER OF SEQ ID NOS: 725
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 209
; LENGTH: 962
; TYPE: DNA
; ORGANISM: Mouse
US-10-152-661-209

Query Match 10.9%; Score 82.6; DB 9; Length 962;
Best Local Similarity 52.5%; Pred. No. 2,2e-12;
Matches 287; Conservative 0; Mismatches 239; Indels 21; Gaps 4;

Qy 112 TGGCAGCGCTTCAAGATGATGCGAGGAATCATCCCGAGCTCGAGAGTACCCCGAG 171

Db 107 TGTTCGCTTTTAAATATGATGCGACAGAAATTCCTTTATTCATGTCGTTAAACCTGTC 166
QY 172 COTCCAGCGAGCTGGAGAACAAAGACATGAACCGGGGGGAGAAACGAGGGGGGCT 231
Db 167 CCGGAC---ACCCAGCAGCAACAGCACCCTGAATCAAGCAGAGATGGAGGAGGAT 223
QY 232 CCCACCAACC---CTTTAGAACCAAGACGTCGAGTCAAGCTGCGGAGCTGAC 288
Db 224 TTCACTAGCAGCTGGAGCTGGATCGAAACAGTCGAGTTCAAGTGGGCTGGAGGAACTGCGG 283
QY 289 TTCAACCCGCTACGTGACCGATGGGCGCTGCGGAGCGCAAGCGGTCACCGAGCTGTG 348
Db 284 TCACCAAAATCATTTTCGAGCGGCGACAGTCACACCTCTGAAGAGACTGTGTG 343
QY 349 TGCTCCGCGCAGTGGCGGCGCGGCGCTGCTGCGCCAAACGCGATCGGCGCGGCGC----- 402
Db 344 TCGCGGGCGAGTGTCTTCCCTCCCTGCGGTCCTTCCCACTGGATCGAGAGAGGCTTACGGA 403
QY 403 -----AAGTGGTGGCACTAGTGGGCGCGACTTCGCTGCATCCCGACCGCTACCGC 456
Db 404 ACAAACTACTGGAGCGGAGGAGCTCTCAGAGTGGGCGGTGTCAACGACAAAGACGCGC 463
QY 457 GCGCAGCGGCTGCAAGCTGTGTCCGCTGTGAGGCGCGCGCGCGCAAGGTGCGC 516
Db 464 ACCCAGAGATCGAGCTGCAAGTGTCAAGACG---CAGCAGCGCGCACCTACAAATACACC 520
QY 517 CTGGTGGCTGTGCAATGCAAGGCGCTCAACCGCTTCCAAACAGTGGAGCTGAG 576
Db 521 GTGTGTCAGGCTGCAAGTGTCAAGAGTACCCGTGACGACCAAGTCCAGCACAAC 580
QY 577 GACTTCGGAGCGGAGCGCTGCGCGCGCAGAGGCGCGGAGCGCGCGCGCGCGCG 636
Db 581 TTTGAAAGCGTGTGCGCGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 640
QY 637 AGCGCCA 643
Db 641 AAATCCA 647

RESULT 9

US-09-866-050A-40
; Sequence 40, Application US/09866050A
; Publication No. US20030040471A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; FILE REFERENCE: 11000.1011c4U
; CURRENT APPLICATION NUMBER: US/09/866,050A
; CURRENT FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 725
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 40
; LENGTH: 962
; TYPE: DNA
; ORGANISM: Mouse
US-09-866-050A-40

Query Match 10.9%; Score 82.6; DB 9; Length 962;
Best Local Similarity 52.5%; Pred. No. 2.2e-12;
Matches 287; Conservative 0; Mismatches 219; Indels 21; Gaps 4;
QY 112 TGGCAGGCGCTTCAAGATGATGCGACGGAATCATCCCGAGCTCGAGAGTACCCCGAG 171
Db 107 TGTTCGCTTTTAAATATGATGCGACAGAAATTCCTTTATTCATGTCGTTAAACCTGTC 166
QY 172 COTCCAGCGAGCTGGAGAACAAAGACATGAACCGGGGGGAGAAACGAGGGGCGCT 231

Db 167 CCGGAC---ACCCAGCAGCAACAGCACCTGTAATCAAGCAGGAAATGAGGAGGAGCAT 223
QY 232 CCCACCAACC---CTTTAGAACCAAGACGTCGTCGAGTACAGTGTGCGGAGCTGAC 288
Db 224 TTCACTAGCAGCTGGAGCTGGATCGAAACAGTCGAGTTCAAGTGGGCTGGAGGAACTGCGG 283
QY 289 TTCAACCCGCTACGTGACCGATGGGCGCTGCGGAGCGCAAGCGGTCACCGAGCTGTG 348
Db 284 TCACCAAAATCATTTTCGAGCGGCGACAGTCACACCTCTGAAGAGACTGTGTG 343
QY 349 TGCTCCGCGCAGTGGCGGCGCGGCGCTGCTGCGCCAAACGCGATGGCGCGGCGC----- 402
Db 344 TCGCGGGCGAGTGTCTTCCCTCCCTGCGGTCCTTCCCACTGGATCGAGAGGCTTACGGA 403
QY 403 -----AAGTGGTGGCACTAGTGGGCGCGACTTCGCTGCATCCCGACCGCTACCGC 456
Db 404 ACAAACTACTGGAGCGGAGGAGCTCTCAGAGTGGGCGGTGTCAACGACAAAGACGCGC 463
QY 457 GCGCAGCGGCTGCAAGCTGTGTCCGCTGTGAGGCGCGCGCGCGCAAGGTGCGC 516
Db 464 ACCCAGAGATCGAGCTGCAAGTGTCAAGACG---CAGCAGCGCGCACCTACAAATACACC 520
QY 517 CTGGTGGCTGTGCAATGCAAGGCGCTCAACCGCTTCCAAACAGTGGAGCTGAG 576
Db 521 GTGTGTCAGGCTGCAAGTGTCAAGAGTACCCGTGACGACCAAGTTCAGCACAAC 580
QY 577 GACTTCGGAGCGGAGCGCTGCGCGCGCAGAGGCGCGGAGCGCGCGCGCGCGCGCG 636
Db 581 TTTGAAAGCGTGTGCGCGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 640
QY 637 AGCGCCA 643
Db 641 AAATCCA 647

RESULT 10

US-09-866-050A-209
; Sequence 209, Application US/09866050A
; Publication No. US20030040471A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: and Methods for Their Use
; FILE REFERENCE: 11000.1011c4U
; CURRENT APPLICATION NUMBER: US/09/866,050A
; CURRENT FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 725
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 209
; LENGTH: 962
; TYPE: DNA
; ORGANISM: Mouse
US-09-866-050A-209

Query Match 10.9%; Score 82.6; DB 9; Length 962;
Best Local Similarity 52.5%; Pred. No. 2.2e-12;
Matches 287; Conservative 0; Mismatches 219; Indels 21; Gaps 4;
QY 112 TGGCAGGCGCTTCAAGATGATGCGACGGAATCATCCCGAGCTCGAGAGTACCCCGAG 171
Db 107 TGTTCGCTTTTAAATATGATGCGACAGAAATTCCTTTATTCATGTCGTTAAACCTGTC 166
QY 172 COTCCAGCGAGCTGGAGAACAAAGACATGAACCGGGGGGAGAAACGAGGGGCGCT 231
Db 167 CCGGAC---ACCCAGCAGCAACAGCACCCTGAATCAAGCAGGAAATGGAGGAGGAT 223
QY 232 CCCACCAACC---CTTTAGAACCAAGACGTCGAGTACAGTGTGCGGAGCTGAC 288

DB 224 TTCAGTAGCACTGGACTGATTCGAAACATGTCAGTTCAGTGGGCTGACAGGAATGCGG 283
 QY 289 TTCACCCGCTACGTAACGATGGGCGGTGCGAGCGCAAGCCGGTCAACCGAGCTGGTG 348
 DB 284 TCACCAAAATACATTTTGGACGGCGAGTGCACAGCATCAGCCCTTGAAGAGCTGGTG 343
 QY 349 TGTCTCCGCGCAGTGGCGCCGCGCGCTCTGTCGCCAACCCATCGGCGCGCG----- 402
 DB 344 TGGCGGGGCGAGTGTGCTTGGCCCTGCGGTGTCTTCCAACTGATTCGAGAGAGCTTACGA 403
 QY 403 -----AAGTGTGGCGCACTAGTGGGCGCGACTTCCTGCTCATTCGCCAGCGCTACCGG 456
 DB 404 ACAAAATACAGAGCGCGAGAGCTCTCAGAGTGGCGGTGTCTCAACGACAAGAGCGCG 463
 QY 457 GCGCAGCGCGTGCAGCTGTGTCTCCGCGGTGTGAGGCGCGCGCGCGCAAGGTGGCG 516
 DB 464 ACCCAGAGATTCAGCTGTGCTGCTCAGAGCG---CAGCAGCGCGCATTAACAAATCACC 520
 QY 517 CTGTGGCTCTGTGCAAGTGCAGAGCGCTTCCAGCTTCCAAACAGTGCAGTCAAG 576
 DB 521 GTGTGTCAGCGCTGCAAGTGCAGAGAGTACACCGTCAAGCAACAGTCAAGCAAC 580
 QY 577 GACTTGGGACCAAGCGCGCTCGCGCGCAGAGAGCGCGCGCGCGCGCGCG 636
 DB 581 TTGAAAGCGTGTGCGCAGCGCAAGCGCGCGCAGCAGCAGAGAGCGCGAGAGCGCAG 640
 QY 637 AGCGCCA 643
 DB 641 AATTCGA 647

RESULT 11

US-09-853-625B-1
 ; Sequence 1, Application US/09853625B
 ; Patent No. US20020049304A1

GENERAL INFORMATION:
 APPLICANT: Hastings, Gregg A. and Adams, Mark D.
 TITLE OF INVENTION: Human CCN-Like Growth Factor
 NUMBER OF SEQUENCES: 20
 CORRESPONDENCE ADDRESS:

ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLIAN,
 CECCHI, STEWART & OLSTEIN
 STREET: 6 BECKER FARM ROAD
 CITY: ROSELAND
 STATE: NEW JERSEY
 COUNTRY: USA
 ZIP: 07068

COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5 INCH DISKETTE
 COMPUTER: IBM PS/2
 OPERATING SYSTEM: MS-DOS
 SOFTWARE: WORD PERFECT 5.1

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/853,625B
 FILING DATE: 14-May-2001
 CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 09/053,587
 FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:
 NAME: MULLINS, J.G.
 REGISTRATION NUMBER: 33,073
 REFERENCE/DOCKET NUMBER: 325800-442

TELECOMMUNICATION INFORMATION:
 TELEPHONE: 201-994-1700
 TELEFAX: 201-994-1744

INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 900 BASE PAIRS
 TYPE: NUCLEIC ACID
 STRANDEDNESS: SINGLE
 TOPOLOGY: LINEAR

MOLECULE TYPE: cDNA

SEQUENCE DESCRIPTION: SEQ ID NO: 1:
 US-09-853-625B-1

Query Match 9.2%; Score 69.6; DB 10; Length 900;
 Best Local Similarity 50.0%; Pred. No. 4.6e-09;
 Matches 239; Conservative 0; Mismatches 224; Indels 15; Gaps 2;

QY 110 GGTGGCAGGCGTTCAAGATGATGCGAAGATATCCCGAGCTGGAGATGCTCCG 169
 DB 176 GCTGTTGGCTTTTAAATATGATGCGAAGATCTTTATTCATATGTGTTAACTTG 235
 QY 170 ACCCTTCAACCGAGCTGAGAGAACAAAGACATGAAACCGGCGGAGAACGAGCGCGC 229
 DB 236 TTCAGACACACCCAGCAGCAACAGACAGTGAATCAAGCCAGAAATGAGAGCGACTT 295
 QY 230 CTCGCCACACCCCTTTTGAACCAAGACGTGTCCGAGTACAGCTGCGCGAGCTGCACT 289
 DB 296 TCAGTAAACATGGAATGGAATCGGAACACTCGGCTTCAAGTGGGTGCGGGAACGCTG 355
 QY 290 TCACCCGCTACGTGACCGATGGGCGGTGCGAGCGCCAGCGCGGTCAACCGAGCTGTGT 349
 DB 356 CCACCAATACATCTCTGATGAGCCAGTGCACACGATACAGCTTGAAGAGCTGTGT 415
 QY 350 GCTCCGCGCAGTGGCGCGCGCGCGCTGTGCCCCAACGCCATCGGCGCGC----- 402
 DB 416 GTGTGGCGAGTGTGCTGCGCGCGCGCGCGCTTCACTGATGAGAGAGCTATGGA 475
 QY 403 -----AAGTGTGGCGAGCTATGAGGCGCGCGCTTCCGCTCATCCCGAGCTTACCGCG 457
 DB 476 CAAAGTACTGGAGCAGAGAGAGCTCCAGAGTGGCGGTGTCAATACAAACCCGTA 535
 QY 458 GCGCAGCGGTGAGCTGTGTCTCCGCGGTGAGAGCGCGCGCGCGCAAGGTGGCGC 517
 DB 536 CCCAAGATACATCAGCTGCAAGTGCAGAGTGG---CAGCAGCGCATTAACAAATCAGAG 592
 QY 518 TGTGGCGCTGTGCAAGTGCAGAGCGCTTACCGCTTCCAAACAGTGCAGAGTCAA 575
 DB 593 TAGTACATGCGCTGCAAGTGCAGAGAGTACACCGGCGCAGCAACAGATCCAGTCA 650

RESULT 12

US-09-745-288-92/c
 ; Sequence 92, Application US/09745288
 ; Patent No. US20010018058A1

GENERAL INFORMATION:
 APPLICANT: Reed, Steven G.
 APPLICANT: Xu, Jiangchun

APPLICANT: Dillon, David C.
 TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND
 TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE

FILE REFERENCE: 210121.446D1
 CURRENT APPLICATION NUMBER: US/09/745,288
 CURRENT FILING DATE: 2000-12-19

NUMBER OF SEQ ID NOS: 101
 SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 92

LENGTH: 1692
 TYPE: DNA

ORGANISM: Homo sapien
 US-09-745-288-92

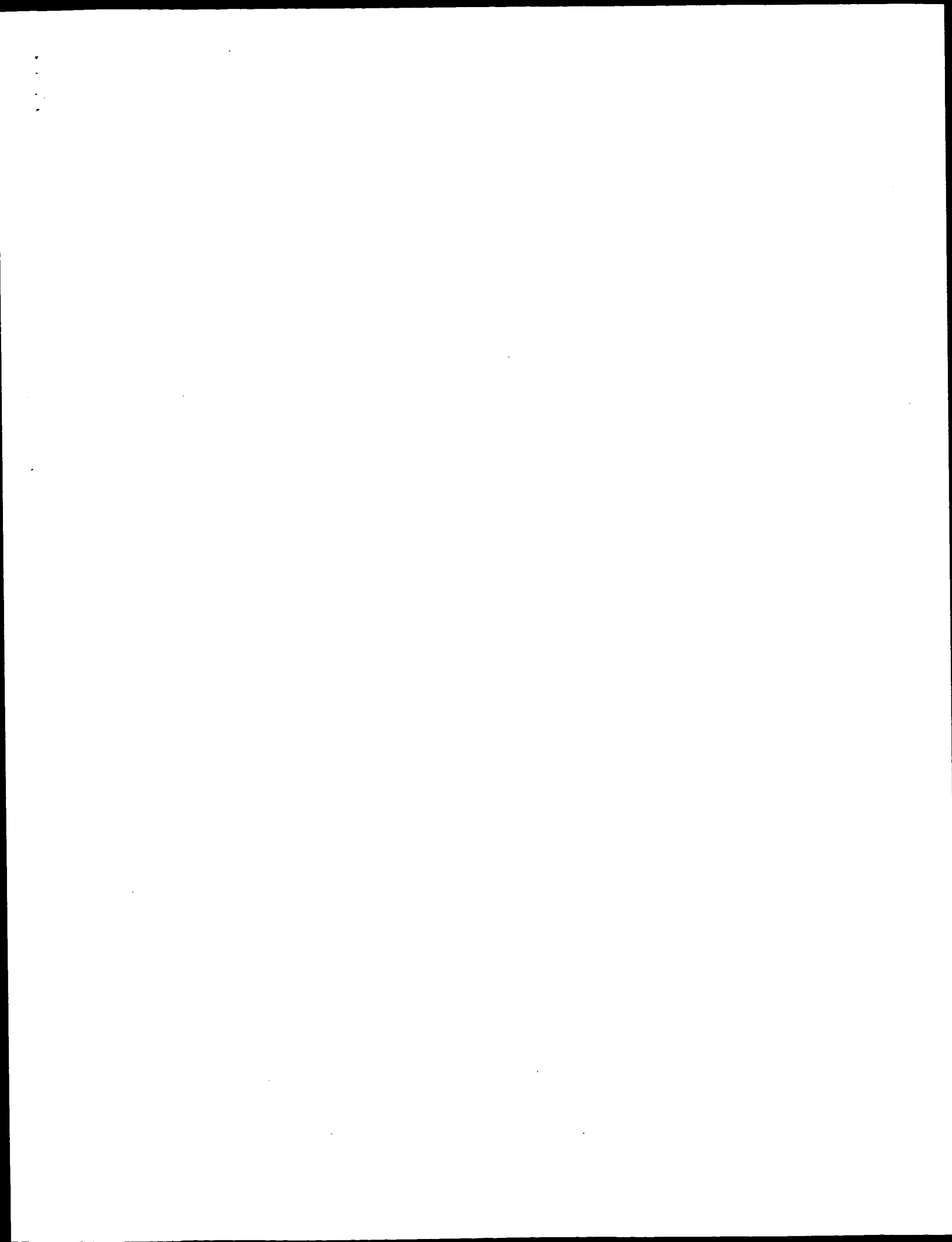
Query Match 9.2%; Score 69.6; DB 10; Length 1692;
 Best Local Similarity 50.0%; Pred. No. 4.5e-09;
 Matches 239; Conservative 0; Mismatches 224; Indels 15; Gaps 2;

QY 110 GGTGGCAGGCGTTCAAGATGATGCGAAGATATCCCGAGCTGGAGATGCTCCG 169
 DB 1582 GCTGTTGGCTTTTAAATATGATGCGAAGATCTTTATTCATATGTGTTAACTTG 1523
 QY 170 AGCTCCACCGAGCTGAGAGAACAAAGACATGAAACCGGCGGAGAGAGCGCGC 229
 DB 1522 TTCAGACACACCCAGCAGCAACAGACAGTGAATCAAGCCAGAAATGAGAGCGACTT 1463

Query Match	6.5%	Score 49;	DB 10;	Length 4098;
Best Local Similarity	44.2%;	Pred. No. 0.00077;		
Matches 202; Conservative	0;	Mismatches 255;	Indels 0;	Gaps 0

	Query Match	Score	DB 107	Length	4098
	Best Local Similarity	44.2%	Pred. No.	0.00077	
	Matches	202	Conservative	0	Mismatches 255
					Indels 0
					Gaps
Qy	268	TTACACTGCCCCGAGACTGCATCTTACCCGGCTACGTCACCGATGGAGCCGTGCGAGGCC	327		
Db	473	TTCGCCCGCCGCGCGCGCGCGCGAGCAGCGCGCCGCGCGCTTCTTCGAAACAGGTC	414		
Qy	328	AAGCCGCTACCCGAGCTGATGTCGCGAGCAGTCGCGCGCGCGCGCGCTGTCTGCCAAC	387		
Db	413	CACCTGAAACGGCTCTTGCTTCGGGTGGGCCCCAAGGATTTGCGCCAGCCGCTCCCGC	354		
Qy	388	GCCATCGCCGCGGCGCAAGTGTGTGCGACTTATGTGGCCCGCATTCGCGTGCATCCCGAC	447		
Db	353	CGCCGCGGAGGCTTCATGCGCGAGCCCGCGCGCGCTCCCGGCTCGCGGACGAGATCTC	284		
Qy	448	CGCTACCGCGCGCAGCGCGGTGACAGTGTGTGTCCCGGTGTAGAGCGCGCGCGCGC	507		
Db	293	CGGACCCGAGGCGACAGCGCGCGCGCGAGTTCCACCCCTGGCTGGCGCAGCGAGCGC	234		
Qy	508	AAGGTGCGCTGTGTGGCTCTGTGCAAGTGCAGCGCGCTTCACCGCTTTCACCAACAGTCG	567		
Db	213	TGACGCGCTTTCGCCGACCCCGGCAAGTCCCGGCGCCCGGAGGAGGCCCGTGGCGCG	174		
Qy	568	GAGCTCAAGGACTTCGGGACCGAGAGCGCGCTCGCGCGGACGAAGGCGCGGAAAGCCCGCGCC	627		
Db	173	CGGCTTCATAGCTGCGCGGACCCCGCGGAACTGCAGAGCCCGCGGACGTCCTCCGCGAG	114		
Qy	628	CGCGCCGCGAGCGCCAAAGCCCAACAGAGCGCGAGCTTGAGAGACGCTTACTTGAAGCCGCGCC	687		
Db	113	CAGCGCTTCACACCCCTGCGCGCGGTGGCGACAGCGCGGACAGACGCTGAGAGTCT	54		
Qy	688	GCGCGCCCTCCCGACCGGCGCGCGCGCGCGCGCGCGCTTGAAC	724		
Db	53	TTTCTCCCGCGCTTGGCGGCGCGCGCTTGAAGCTT	17		

Search completed: March 29, 2003, 01:33:21
Job time : 98.8624 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 29, 2003, 00:44:05 ; Search time 2465.25 Seconds

(without alignments)
8960.144 Million cell updates/sec

Title: US-09-867-274-1

Perfect score: 759

Sequence: 1 tctggaaggtggtggtcc.....ctgtccctcgtggttt 759

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 2054640 seqs, 14551402878 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4139280

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

```
Database :
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_stb:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_or:*
21: em_ov:*
22: em_ov:*
23: em_ov:*
24: em_pl:*
25: em_pl:*
26: em_ro:*
27: em_stb:*
28: em_un:*
29: em_un:*
30: em_hcg_hum:*
31: em_hcg_inv:*
32: em_hcg_inv:*
33: em_hcg_inv:*
34: em_hcg_inv:*
35: em_hcg_inv:*
36: em_hcg_inv:*
37: em_hcg_inv:*
38: em_hcg_inv:*
39: em_hcg_inv:*
40: em_hcg_inv:*
41: em_hcg_inv:*
```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	759	100.0	759	6	AX323453
2	759	100.0	2296	9	AX323454
3	759	100.0	2232	9	AX323455
4	757	99.7	2329	6	AX056687
5	735	96.8	2271	6	AX342535
6	505	66.5	7099	6	AX342537
7	505	66.5	21501	9	AX326736
8	505	66.5	80117	9	AC055813
9	505	66.5	94752	9	AC003098
10	220	29.0	177744	2	AC073954
11	110	14.5	642	9	AF326742
12	83	10.9	532	4	AF326738
13	81	10.7	93790	9	AF397423
14	69	9.1	636	6	AX323455
15	69	9.1	638	10	AF326740
16	69	9.1	81806	10	AF326737
17	69	9.1	110000	2	AC068782-2
18	69	9.1	205277	2	AC012296
19	69	9.1	208135	10	AC068807
20	68	9.0	674	10	AF326741
21	68	9.0	101804	2	AC098160
22	68	9.0	104898	2	AC121721
23	40	5.3	40	6	AX056701
24	35	4.6	198508	10	AL591145
25	27	3.6	27	6	AX056700
26	27	3.6	27	6	AX323473
27	26	3.4	26	6	AX323460
28	24	3.2	24	6	AX323459
29	23	3.0	23	6	AX323474
30	23	3.0	23	6	AX342540
31	22	2.9	51575	2	AC023810
32	22	2.9	170472	2	AC126215
33	22	2.9	194746	2	AC123544
34	21	2.8	21	6	AX056699
35	21	2.8	41	6	AX323472
36	21	2.8	1862	9	AK093966
37	21	2.8	2266	6	AR073028
38	21	2.8	2266	9	HIMELK1A
39	21	2.8	2695	9	AB016193
40	21	2.8	3701	9	AF080617
41	21	2.8	4935	12	AF050499
42	21	2.8	4935	12	AF063584
43	21	2.8	6649	9	AF080618
44	21	2.8	6847	9	HOELK2
45	21	2.8	10558	9	AB016195

ALIGNMENTS

```
RESULT 1
AX323453 LOCUS AX323453 759 bp DNA linear PAT 07-JAN-2002
Sequence 1 from Patent WO0192308.
AX323453 ACCESSION AX323453
VERSION AX323453.1 GI:18094216
KEYWORDS
SOURCE
ORGANISM human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
AUTHORS Paszty,C.J. and Gao,Y.
TITLE Cysteine-knot polypeptides: cloaked-2 molecules and uses thereof
JOURNAL Patent: WO 0192308-A 1 06-DEC-2001;
```


Qy	421	GGGCGCGACTTCGCGTGCATCCCGCAACCGCTACGCGGCGACCGGTGTAAGCTGCTGT	480
Db	422	GGGCGCGACTTCGCGTGCATCCCGCAACCGCTACGCGGCGACCGGTGTAAGCTGCTGT	481
Qy	481	CCCGGTGCTGAGGCGCCGCGCCGCGCAAGTGGCGCTGTGGTGCCTCTGTGCAGTGCAG	540
Db	482	CCCGGTGCTGAGGCGCGCCGCGCGCAAGTGGCGCTGTGGTGCCTCTGTGCAGTGCAG	541
Qy	541	CGCTCACCCGCTTCCACACACGTCGAGCTCAAGAGCTTCGGGACCGAGGCGCTCGG	600
Db	542	CGCTCACCCGCTTCCACACACGTCGAGCTCAAGAGCTTCGGGACCGAGGCGCTCGG	601
Qy	601	CCGCGAAGAGGCGCGGAACCGCGGCGCCGCGCGGAGCGCCAAAGCCACACAGGCGGAG	660
Db	602	CCGCGAAGAGGCGCGGAACCGCGGCGCCGCGCGGAGCGCCAAAGCCACACAGGCGGAG	661
Qy	661	CTGGAGAACGCGCTTACTAGAGCCGCGCCGCGCCCTCCCAACGGAGGGGAGCCCGCGGCTT	720
Db	662	CTGGAGAACGCGCTTACTAGAGCCGCGCCGCGCCCTCCCAACGGAGGGGAGCCCGCGGCTT	721
Qy	721	GAACCGCGCGCCACATTTCTGTCTCTGCGCGGTGGTTT	759
Db	722	GAACCGCGCGCCACATTTCTGTCTCTGCGCGGTGGTTT	760

RESULT 3	AF326739	2323 bp	mRNA	linear	PRI 28-FEB-2003
LOCUS	AF326739				
DEFINITION	Homo sapiens sclerostin				
ACCESSION	AF326739				
VERSION	AF326739.1				
KEYWORDS	GI:13161019				
SOURCE	Homo sapiens.				

SOURCE
ORGANISM

REFERENCE AUTHORS

TITLE	JOURNAL
Bone dysplasia scleroosteosis results from loss of the SOST gene product, a novel cysteine knot-containing protein	Am. J. Hum. Genet. 68 (3), 577-589 (2001)

REFERENCE
AUTHORS

TITLE Direct Submission
Submitted (07-DEC-2000) Genomics, Celltech Chiroscience Inc., 1631
JOURNAL *Genomics*

SOURCE

5'UTR
CDS

3' UTR	BASE COUNT
1	1
2	2
3	3
4	4
5	5
6	6
7	7
8	8
9	9
10	10
11	11
12	12
13	13
14	14
15	15
16	16
17	17
18	18
19	19
20	20
21	21
22	22
23	23
24	24
25	25
26	26
27	27
28	28
29	29
30	30
31	31
32	32
33	33
34	34
35	35
36	36
37	37
38	38
39	39
40	40
41	41
42	42
43	43
44	44
45	45
46	46
47	47
48	48
49	49
50	50
51	51
52	52
53	53
54	54
55	55
56	56
57	57
58	58
59	59
60	60
61	61
62	62
63	63
64	64
65	65
66	66
67	67
68	68
69	69
70	70
71	71
72	72
73	73
74	74
75	75
76	76
77	77
78	78
79	79
80	80
81	81
82	82
83	83
84	84
85	85
86	86
87	87
88	88
89	89
90	90
91	91
92	92
93	93
94	94
95	95
96	96
97	97
98	98
99	99
100	100

ORIGIN

Query Match	100.0%	Score 759;	DB 9;	Length 2323;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 759; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;

QY	1	TACTGGAAGGTGGCGTCGCTCTCTGCGCTGGTACCAATGCAAGCTCCCACTGGCCCTGTGT	60
Db	12	TACTGGAAGGTGGCGTCGCTCTCTGCGCTGGTACCAATGCAAGCTCCCACTGGCCCTGTGT	71
QY	61	CTCGTCTGCTGCTGTGTACACAACAGCTTTCCGTGTGTGTGTGAGAGGCGCAGAGGGTGTGCAGGCG	120
Db	72	CTCGTCTGCTGCTGTGTGTACACAAGCTTTCCGTGTGTGTGTGAGAGGCGCAGAGGGTGTGCAGGCG	131
QY	121	TTCAAGATGTATGCCAGGAAATCAATCCCGAGCTCGGAGAGTACCCGAGGCTCCACCG	180
Db	132	TTCAAGATGTATGCCAGGAAATCAATCCCGAGCTCGGAGAGTACCCGAGGCTCCACCG	191
QY	181	GAGCTGAGAAACAACAAGCATGAAACGGGCGGAGAAACGAGGGCGGCTTCCACACAC	240
Db	192	GAGCTGAGAAACAACAAGCATGAAACGGGCGGAGAAACGAGGGCGGCTTCCACACAC	251
QY	241	CCCTTTGAGACCAAGAAGCGTGTCCGAGTACAGCTGCGGAGCTGCACCTTCAACCCGCTAC	300
Db	252	CCCTTTGAGACCAAGAAGCGTGTCCGAGTACAGCTGCGGAGCTGCACCTTCAACCCGCTAC	311
QY	301	GTGACCGATGGGCGCGTGGCCGACGCGCCAAGCCGCTCACCCGAGCTGTGTGCTCCGGCCAG	360
Db	312	GTGACCGATGGGCGCGTGGCCGACGCGCCAAGCCGCTCACCCGAGCTGTGTGCTCCGGCCAG	371
QY	361	TGCGGCGCGGGCGCGCTGCTGTGCCAAAGCCATTCGGCGCGCGGCAATGGTGGGAGCCTTGT	420
Db	372	TGCGGCGCGGGCGCGCTGCTGTGCCAAAGCCATTCGGCGCGCGGCAATGGTGGGAGCCTTGT	431
QY	421	GGGCGCGCATTCGCTGCGCTGCATCCCGACCGCTTACCGCGCGCAGCGCGTGCATGCTGTGT	480
Db	432	GGGCGCGCATTCGCTGCGCTGCATCCCGACCGCTTACCGCGCGCAGCGCGTGCATGCTGTGT	491
QY	481	CCCGGTGTGAGGGCGCGCGCGCGCGCAAGGTGCGCTGTGTGAGCTTGGTGCATGTGCAAG	540
Db	492	CCCGGTGTGAGGGCGCGCGCGCGCGCAAGGTGCGCTGTGTGAGCTTGGTGCATGTGCAAG	551
QY	541	GCGCTCACCGGCTTCCACAACAGTCGAGAGCCAAAGGACTTCGGGACCGAGGCGGCTCGG	600
Db	552	GCGCTCACCGGCTTCCACAACAGTCGAGAGCTTCGGGAGCTTCGGGACCGAGGCGGCTCGG	611
QY	601	CCGCGAAGAGGCGCGGAAGCCGCGGCGCGCGCGCCGCGGAGCGGCCAAAGCCACAGGCGGAG	660
Db	612	CCGCGAAGAGGCGCGGAAGCCGCGGCGCGCGCGCCGCGGAGCGGCCAAAGCCACAGGCGGAG	671
QY	661	CTGAGAAACGCTTACTAGAGCCCGCGCGCGCTTCCCAACCGGCGGCGCGCCCGGCTT	720
Db	672	CTGAGAAACGCTTACTAGAGCCCGCGCGCGCTTCCCAACCGGCGGCGCGCCCGGCTT	731
QY	721	GAACCCGCGGCCCACTTCTGTCTCTCTGAGCGCGTGT	759
Db	732	GAACCCGCGGCCCACTTCTGTCTCTCTGAGCGCGTGT	770

RESULT 4	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE
AX056687	AX056687	Sequence 19 from Patent WO0075317.	AX056687	AX056687.1	GI:12309667	human.

REFERENCE
AUTHORS

1 (bases 1 to 2329)
Botstein, D.A., Goddard, A., Gurney, A.L., Smith, V., Watanabe, C.K. and Wood, W.I.

Db 721 CTCTGGCGGTGTT 725

RESULT 6
LOCUS AX342537/c
DEFINITION Sequence 3 from Patent WO0198491.
ACCESSION AX342537
VERSION AX342537.1 GI:18151965
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS Bailemans, W., Ebeling, M., Foenzler, D., Patel, N., van Hul, W. and Vickers, B.H.
TITLE Osteolevin gene polymorphisms
JOURNAL Patent: WO 0198491-A 3 27-DEC-2001; F. HOFMANN-LA ROCHE AG (CH); UNIVERSITAIRE INSTELLING ANTWERPEN (BE)

FEATURES
source Location/Qualifiers
1..7099
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 1530 a 1944 c 1928 g 1697 t

ORIGIN

Query Match 66.5%; Score 505; DB 6; Length 7099;
Best Local Similarity 100.0%; Pred. No. 6e-242;
Matches 505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 255 AGAGGTGTCCGAGTACAGCTGCGCGAGCTGCACTTACCCGCTACCGATGGGCC 314
Db 2343 AGAGGTGTCCGAGTACAGCTGCGCGAGCTGCACTTACCCGCTACCGATGGGCC 2284

QY 315 GTGCCGAGCGCCAGCGCGTCACTGCTGCTCCGCGCAAGTCCGCGCGCG 374
Db 2283 GTGCCGAGCGCCAGCGCGTCACTGCTGCTCCGCGCAAGTCCGCGCGCG 2224

QY 375 CTTGCTCCCAACCGCATCGCGCGCGCAAGTGTGGGCACTTATGGGCGCG 434
Db 2223 CTTGCTCCCAACCGCATCGCGCGCGCAAGTGTGGGCACTTATGGGCGCG 2164

QY 435 CTGCAATCCCGACCGCTACCGCGCGCGCAAGTGTGGGCACTTATGGGCGCG 494
Db 2163 CTGCAATCCCGACCGCTACCGCGCGCGCAAGTGTGGGCACTTATGGGCGCG 2104

QY 495 GCGCGCGCGCGCAAGTGTGGGCACTTATGGGCGCGCGCGCGCGCGCGCG 554
Db 2103 GCGCGCGCGCGCAAGTGTGGGCACTTATGGGCGCGCGCGCGCGCGCGCG 2044

QY 555 CCAACAACAGTGGAGTCAAGGACTTGGGAGCGGAGCGCGCGCGCGCGCGCG 614
Db 2043 CCAACAACAGTGGAGTCAAGGACTTGGGAGCGGAGCGCGCGCGCGCGCGCG 1984

QY 615 GAAAGCG 674
Db 1983 GAAAGCG 1924

QY 675 CTGAGAGCG 734
Db 1923 CTGAGAGCG 1864

QY 735 CATTTCTGTCTCTGCGCGGTGTT 759
Db 1863 CATTTCTGTCTCTGCGCGGTGTT 1839

RESULT 7
LOCUS AF326736
DEFINITION Homo sapiens sclerostin gene, complete cds.

ACCESSION AF326736
VERSION AF326736.1 GI:13161010
KEYWORDS
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS Brunkow, M.E., Gardner, J.C., Van Ness, J., Paep, B.W., Kovacevich, B.R., Prohl, S., Skonier, J.E., Zhao, L., Sabo, P.J., Fu, Y.H., Alisch, R.S., Gillet, L., Colbert, T., Tacconi, P., Galas, D., Hamersma, H., Beighton, P. and Mulligan, J.T.
TITLE Bone dysplasia osteostosis results from loss of the SOST gene product, a novel cysteine knot-containing protein
JOURNAL Am. J. Hum. Genet. 68 (3), 577-589 (2001)
MEDLINE 21090529
PUBMED 11179006

REFERENCE 2 (bases 1 to 21501)
AUTHORS Brunkow, M.E., Gardner, J.C., Van Ness, J., Paep, B.W., Kovacevich, B.R., Prohl, S., Skonier, J.E., Zhao, L., Sabo, P.J., Fu, Y.H., Alisch, R.S., Gillet, L., Colbert, T., Tacconi, P., Galas, D., Hamersma, H., Beighton, P. and Mulligan, J.T.
TITLE Direct Submission
JOURNAL Submitted (07-DEC-2000) Genomics, Celltech Chiroscience Inc., 1631 220th St. SE, Bothell, WA 98021, USA

FEATURES
source Location/Qualifiers
1..21501
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="17"
/map="17q21"
join(111961..12180,14939..15360)
/product="sclerostin"
join(111961..12180,14939..15360)
/codon_start=1
/product="sclerostin"
/protein_id="AAK13451.1"
/translation="MDPLALCLVLTAFRVVGGQWQAFNDATTEIIPEIGEP
BPPELENNKTMRAENGRRPDRHPEFEDKSEVSCPLHETRYVTGPGSARVTE
LVSCQCPRALIPRAITGRKMPKSGDFPCIDPRYDAQVQVLCPEGELPRKRVK
LVASCKRLTRFNNQSELDKFGTEAARPRGRPRPARSAKANQAELENAY"

BASE COUNT 5546 a 5259 c 5419 g 5277 t

ORIGIN

Query Match 66.5%; Score 505; DB 9; Length 21501;
Best Local Similarity 100.0%; Pred. No. 5.9e-242;
Matches 505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 255 AGAGGTGTCCGAGTACAGCTGCGCGAGCTGCACTTACCCGCTACCGATGGGCC 314
Db 14937 AGAGGTGTCCGAGTACAGCTGCGCGAGCTGCACTTACCCGCTACCGATGGGCC 14996

QY 315 GTGCCGAGCGCCAGCGCGTCACTGCTGCTCCGCGCAAGTCCGCGCGCG 374
Db 14997 GTGCCGAGCGCCAGCGCGTCACTGCTGCTCCGCGCAAGTCCGCGCGCG 15056

QY 375 CTTGCTCCCAACCGCATCGCGCGCGCGCAAGTGTGGGCACTTATGGGCGCG 434
Db 15057 CTTGCTCCCAACCGCATCGCGCGCGCGCAAGTGTGGGCACTTATGGGCGCG 15116

QY 435 CTGCAATCCCGACCGCTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 494
Db 15117 CTGCAATCCCGACCGCTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 15176

QY 495 GCG 554
Db 15177 GCG 15236

QY 555 CCAACAACAGTGGAGTCAAGGACTTGGGAGCGGAGCGCGCGCGCGCGCGCG 614
Db 15237 CCAACAACAGTGGAGTCAAGGACTTGGGAGCGGAGCGCGCGCGCGCGCGCG 15296


```

repeat_region complement (2492, .2787)
                /rpt_family="AluSg"
repeat_region complement (2796, .2902)
                /rpt_family="MIR"
repeat_region complement (2903, .3165)
                /rpt_family="AluSg"
repeat_region complement (3166, .3193)
                /rpt_family="MIR"
repeat_region complement (3572, .3644)
                /rpt_family="MIR"
repeat_region complement (4524, .4583)
                /rpt_family="MIR3"
repeat_region complement (4755, .4962)
                /rpt_family="MIR"
repeat_region complement (6397, .6738)
                /rpt_family="L2"
repeat_region complement (6766, .6875)
                /rpt_family="MLT1H"
repeat_region complement (7202, .7270)
                /rpt_family="L2"
repeat_region complement (7739, .7825)
                /rpt_family="L1MC4"
repeat_region complement (7826, .8122)
                /rpt_family="AluSc"
repeat_region complement (8123, .8268)
                /rpt_family="L1MC4"
repeat_region complement (8269, .8568)
                /rpt_family="AluJo"
repeat_region complement (8569, .8631)
                /rpt_family="L1MC4"
repeat_region complement (9151, .9305)
                /rpt_family="MIR"
repeat_region complement (9365, .9390)
                /rpt_family="AT rich"
repeat_region complement (9405, .9708)
                /rpt_family="AluY"
repeat_region complement (9709, .9969)
                /rpt_family="AluJo"
repeat_region complement (10197, .10458)
                /rpt_family="MLT1K"
repeat_region complement (10758, .10895)
                /rpt_family="MIR"
repeat_region complement (12676, .12773)
                /rpt_family="MIR"
repeat_region complement (12784, .12997)
                /rpt_family="MIR"
repeat_region complement (13220, .13513)
                /rpt_family="AluJo"
repeat_region complement (13627, .13787)
                /rpt_family="L2"
repeat_region complement (13821, .13953)
                /rpt_family="MIR"
repeat_region complement (14746, .14899)
                /rpt_family="MIR"
repeat_region complement (14947, .15144)
                /rpt_family="MIR"
repeat_region complement (15520, .15668)
                /rpt_family="AluJo"
repeat_region complement (15665, .15960)
                /rpt_family="AluSx"
repeat_region complement (15961, .16128)
                /rpt_family="AluJo"
repeat_region complement (16208, .16400)
                /rpt_family="MIR"
repeat_region complement (16454, .16586)
                /rpt_family="L2"
repeat_region complement (16707, .16783)
                /rpt_family="MIR"
repeat_region complement (16852, .16909)

```

Query Match 66.5%; Score 505; DB 9; Length 80117;
 Best Local Similarity 100.0%; Fred. No. 5.8e-242;
 Matches 505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 255 AGACGTCGAGTACAGCTGCGGAGCTGCACTTCAACCCGCTACGTGACGATGGCC 314
    |||||
Db 63216 AGACGTCGAGTACAGCTGCGGAGCTGCACTTCAACCCGCTACGTGACGATGGCC 63157
QY 315 GTGCCGACGCGCAAGCCGCGTACCGAGCTGTGTCTCCGCGCATGCGCGCGCG 374
    |||||
Db 63156 GTGCCGACGCGCAAGCCGCGTACCGAGCTGTGTCTCCGCGCATGCGCGCGCG 63097
QY 375 CCTGCTGCCCAACGCCATCTGCGCGCGGCAAGTGTGGGCACTTAATGGCCGACTTCCG 434
    |||||
Db 63096 CCTGCTGCCCAACGCCATCTGCGCGCGGCAAGTGTGGGCACTTAATGGCCGACTTCCG 63037
QY 435 CTGCATCCCGGACCGCTACCGCGCGGCGTGCAGCTGCTGTCCCGGTGGAGGC 494
    |||||
Db 63036 CTGCATCCCGGACCGCTACCGCGCGGCGTGCAGCTGCTGTCCCGGTGGAGGC 62977
QY 495 GCGCGCGCGCGCAAGTGTGCGCTGTGTGCGCTGTGCAAGTGCAGCGCTTCAACCGGCTT 554
    |||||
Db 62976 GCGCGCGCGCGCAAGTGTGCGCTGTGTGCGCTGTGCAAGTGCAGCGCTTCAACCGGCTT 62917
QY 555 CCACACCAAGTGTGAGCTTCAAGACTTGGGACCGAGCGCGTGTGGCGGAGAGGCGCG 614
    |||||
Db 62916 CCACACCAAGTGTGAGCTTCAAGACTTGGGACCGAGCGCGTGTGGCGGAGAGGCGCG 62857
QY 615 GAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTA 674
    |||||
Db 62856 GAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTA 62797
QY 675 CTGAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTA 734
    |||||
Db 62796 CTGAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTA 62737
QY 735 CATTTCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 759
    |||||
Db 62736 CATTTCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 62712

```

RESULT 9

AC003098/c 94752 bp DNA linear PRI 29-JAN-1998
 LOCUS Homo sapiens chromosome 17, clone HRPc905N1, complete sequence.
 AC003098
 AC003098
 VERSION
 KEYWORDS
 HGNC
 SOURCE
 ORGANISM

Homo sapiens.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 AUTHORS
 Birren, B., Fasmann, K., McKernan, K., Nusbaum, C., Richardson, P. and
 Lander, E.

TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 Homo sapiens chromosome 17, clone HRPc905N1
 Unpublished
 2 (bases 1 to 94752)
 Richardson, P., Lander, E., McKernan, K., Munro, C., Nusbaum, C.,
 Cooke, P., Daly, M.J., Devon, K., Dewar, K., Dukette, B., Forrest, C.,
 Gage, D., Gensheimer, S., Geratger, K., Gilmarin, T., Hagos, B.,
 Halphen, I., Harris, K., Howland, J.C., Huang, J., Hui, L., Jacotot, L.,
 Kirby, A., Lane, M., Mackenzie, J., Marquis, N., McDermott, J.,
 Molla, M., Morrow, J., Nachman, A., Naylor, J., O'Connor, T., Olotu, A.,
 Peterson, K., Roberts, D., Rollins, G., Sarnik, A., Shiu, P., Shyam, R.,
 Stillwell, J., Stone, C., Strickland, C., Sydney, K., Tang, L.,
 Zemtseva, I., and Zody, M.

TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 Direct Submission
 Submitted (10-NOV-1997) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 94752)
 Birren, B., Fasmann, K., McKernan, K., Nusbaum, C., Richardson, P.,
 Lander, E., Allen, A., Baker, J., Baldwin, J., Barna, N., Beckert, R.,
 Boutwell, C., Byrne, S., Cantu, C., Castle, A., Cooke, P., Daly, M.J.,
 DePaeye, E., Devon, K., Dewar, K., Dukette, B., Etemadi, S.,

Gerrelley, P., Forrest, C., Gage, D., Gardyna, S., Gensheimer, S.,
Gierulick, K., Gilmartin, T., Gray, D., Hagos, B., Harris, K.,
Horton, L., Howland, J. C., Hul, L., Jacobot, L., Linton, L.,
MacKenzie, U., Marquis, N., McEwan, P., McGuirk, A., Meldrum, J.,
Molla, M., Morris, W., Morrow, J., Nachman, A., Naylor, J., O'Connor, T.,
Pavlin, B., Peterson, K., Ranganath, S., Riley, R., Roberts, D.,
Rollins, G., Rossello, R., Roy, A., Shyam, R., Sochoo, S., Stilwell, J.,
Stone, C., Strickland, C., Sydney, K., Tang, L., Vassiliev, H., Vo, A.,
Wagner, A., Wheeler, J., Wu, Y., Ye, W. J., Zemseva, I., Zhao, J. and
Zody, M.

TITLE Direct Submission
JOURNAL Submitted (29-JAN-1998) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT On Jan 29, 1998 this sequence version replaced gi:2665504. The Staden databases, finishing information, and all chromatographic files used in the assembly of this clone are available from our anonymous ftp site.

All repeats were identified using RepeatMasker: Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>.

FEATURES

source

```

source
1. .94752
/organism="Homo sapiens"
/db xref="taxon:9606"
/chromosome="17"
/map="17"
/clone="HPRC905N1"
/clone_lib="Peter de Jong/ human PAC library"
repeat_region
86. .439
/rpt_family="LINE2"
complement(972. .1262)
/rpt_family="AluXx"
complement(1884. .1912)
/rpt_family="(CA)n"
1912. .1959
/rpt_family="(GA)n"
2633. .2681
/rpt_family="MIR"
2701. .3002
/rpt_family="AluXx"
complement(3874. .4008)
/rpt_family="LINE2"
4513. .4543
/rpt_family="AT_rich"
complement(5202. .5237)
/rpt_family="(GA)n"
complement(5310. .5406)
/rpt_family="MIR"
complement(5829. .5868)
/rpt_family="GC_rich"
6485. .6592
/rpt_family="MIR"
complement(6973. .7029)
/rpt_family="MIR"
complement(10402. .10426)
/rpt_family="GC_rich"
complement(11421. .11496)
/rpt_family="LINE"
complement(11505. .11804)
/rpt_family="AluXx"
11811. .11879
/rpt_family="AT_rich"
11880. .12093
/rpt_family="AluDb"
complement(12104. .12476)
/rpt_family="LINE2"
12476. .12773
/rpt_family="AluXx"
complement(12775. .12917)
/rpt_family="LINE2"
complement(13088. .13510)
/rpt_family="MULTIC"

```

repeat_region	/rpt_family="A1USx"	complement(13547, .13847)
repeat_region	/rpt_family="L1MCI"	complement(14110, .1476)
repeat_region	/rpt_family="A1USx"	complement(14777, .15077)
repeat_region	/rpt_family="A1USx"	complement(15079, .15282)
repeat_region	/rpt_family="L1MC1"	15954, .16024
repeat_region	/rpt_family="AT_rich"	18031, .18178
repeat_region	/rpt_family="WMR"	20398, .20769
repeat_region	/rpt_family="LINE2"	20893, .21100
repeat_region	/rpt_family="AlUSg"	21307, .21952
repeat_region	/rpt_family="MT2CB"	21974, .22115
repeat_region	/rpt_family="WMR"	22238, .22540
repeat_region	/rpt_family="AlUSx"	22577, .22633
repeat_region	/rpt_family="LINE2"	complement(22199, .23270)
repeat_region	/rpt_family="FLAM_A"	24194, .24324
repeat_region	/rpt_family="AlUSp"	24329, .24643
repeat_region	/rpt_family="LINE2"	25542, .25698
repeat_region	/rpt_family="LINE2"	25805, .26001
repeat_region	/rpt_family="WMR"	26060, .26174
repeat_region	/rpt_family="LINE2"	26343, .26649
repeat_region	/rpt_family="AlUSx"	complement(27491, .27783)
repeat_region	/rpt_family="AlUSq"	27862, .27906
repeat_region	/rpt_family="LINE2"	28068, .28114
repeat_region	/rpt_family="WMR"	complement(28329, .28360)
repeat_region	/rpt_family="AT_rich"	28371, .28652
repeat_region	/rpt_family="AlUSg"	complement(29499, .29604)
repeat_region	/rpt_family="GC_rich"	30074, .30181
repeat_region	/rpt_family="WMR"	complement(30181, .30249)
repeat_region	/rpt_family="LINE2"	30444, .30517
repeat_region	/rpt_family="WMR"	complement(31855, .32135)
repeat_region	/rpt_family="AlUSg"	complement(32675, .33003)
repeat_region	/rpt_family="MLT1AI"	complement(33042, .33082)
repeat_region	/rpt_family="LINE2"	33363, .33459
repeat_region	/rpt_family="MER5B"	complement(33475, .35524)
repeat_region	/rpt_family="MLT2D"	35495, .35591


```

repeat_region      /rpt_family="(GA)n"
                    complement(35561..35673)
repeat_region      /rpt_family="MLT2D"
                    complement(35678..35935)
repeat_region      /rpt_family="MER66B"
                    complement(35936..36233)
repeat_region      /rpt_family="AluSg"
                    complement(36236..36449)
repeat_region      /rpt_family="MER66B"
                    complement(36450..36867)
repeat_region      /rpt_family="MLT2D"
                    complement(36876..37460)
repeat_region      /rpt_family="L1MD2"
                    complement(37477..37777)
repeat_region      /rpt_family="AluSx"
                    complement(37804..37948)
repeat_region      /rpt_family="L1"
                    complement(38130..38175)
repeat_region      /rpt_family="L1MD1"
                    complement(38175..38253)
repeat_region      /rpt_family="L1MD2"

```

```

Query Match      66.5%: Score 505; DB 9; Length 94752;
Best Local Similarity 100.0%; Pred. No. 5 Be-242;
Matches 505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 255 AGAGCTGTCCAGTACAGCTCCCGAGCTGCACTTCAACCGCTACGTCACCGATGGGCC 314
Db 6293 AGAGCTGTCCAGTACAGCTCCCGAGCTGCACTTCAACCGCTACGTCACCGATGGGCC 6234
Qy 315 GTGCGCGAGCGCCAGAGCGGTCACCGAGCTGCTGCTCCGCGCAAGTCCGCGCGCG 374
Db 6233 GTGCGCGAGCGCCAGAGCGGTCACCGAGCTGCTGCTCCGCGCAAGTCCGCGCGCG 6174
Qy 375 CTTGCTGCGCCAGCCAGTCGCGCCGCGCAAGTGTGCGCACTAGTGGCCCGACCTTCCG 434
Db 6173 CTTGCTGCGCCAGCCAGTCGCGCCGCGCAAGTGTGCGCACTAGTGGCCCGACCTTCCG 6114
Qy 435 CTGATATCCCGACCGGTCACCGCGCGCGCGCTGTCAGCTCTGTGCTCCGCGTGGAGGC 494
Db 6113 CTGATATCCCGACCGGTCACCGCGCGCGCTGTCAGCTCTGTGCTCCGCGTGGAGGC 6054
Qy 495 GCCCGCGCGCGCGCAAGGTGCGCTGTGCGCTGTGCAAGTGTGCAAGCGCTCAACCGCTT 554
Db 6053 GCCCGCGCGCGCGCAAGGTGCGCTGTGCGCTGTGCAAGTGTGCAAGCGCTCAACCGCTT 5994
Qy 555 CCACACCAAGTCGAGCTCAAGGACCTTGGGACCGAGCGCGCTCGCCAGAGGCGCG 614
Db 5993 CCACACCAAGTCGAGCTCAAGGACCTTGGGACCGAGCGCGCTCGCCAGAGGCGCG 5934
Qy 615 GAAGCGCGCGCGCGCGCGCGCGCGCAAGCAACCAAGCGCGAGCTGAGAGCGCTTA 674
Db 5933 GAAGCGCGCGCGCGCGCGCGCGCGCAAGCAACCAAGCGCGAGCTGAGAGCGCTTA 5874
Qy 675 CTAGAGCGCGCGCGCGCGCGCTTCCCAACCGCGCGCGCGCGCGCTGAAACCGCGCCCA 734
Db 5873 CTAGAGCGCGCGCGCGCGCGCTTCCCAACCGCGCGCGCGCGCGCTGAAACCGCGCCCA 5814
Qy 735 CATTTCTGTCTCTCTGCGCGCTTT 759
Db 5813 CATTTCTGTCTCTCTGCGCGCTTT 5789

```

```

RESULT 10
AC073954      177744 bp      DNA      linear      HTG 17-NOV-2000
LOCUS        AC073954
DEFINITION   Homo sapiens chromosome 17 clone RP11-147C10, WORKING DRAFT
ACCESSION    AC073954
VERSION      AC073954.3   GI:10998993
KEYWORDS     HTG; HTGS PHASE1; HTGS_DRAFT.
SOURCE       Homo sapiens.
ORGANISM     Homo sapiens

```

REFERENCE AUTHORS TITLE JOURNAL REFERENCE AUTHORS TITLE JOURNAL

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 177744)
Waterston, R.H.
The sequence of Homo sapiens clone
Unpublished
2 (bases 1 to 177744)
Waterston, R.H.
Direct Submission
Submitted (07-JUL-2000) Genome Sequencing Center, Washington University School of Medicine, 444 Forest Park Parkway, St. Louis, MO 63108, USA
On Oct 25, 2000 this sequence version replaced gi:186649.

Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
Project Information
Center project name: H.NH0147C10

Summary Statistics

Sequencing vector: M13; 100%
Sequencing vector: plasmid; 0%
Chemistry: Dye-primer ET; 100% of reads
Chemistry: Dye-terminator Big Dye; 0% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 155679 bases at least Q40
Consensus quality: 163329 bases at least Q30
Insert size: 185000; agarose-fp
Insert size: 183559; sum-of-contigs
Quality coverage: 3.29 in Q20 bases; agarose-fp
Quality coverage: 3.43 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 34 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence, as soon as it is available and the accession number will be preserved.

1	1247:	contig of 1247 bp in length
1248	1347:	gap of unknown length
1348	2762:	contig of 1415 bp in length
2763	2862:	gap of unknown length
2863	4773:	contig of 1911 bp in length
4774	4873:	gap of unknown length
4874	6326:	contig of 1453 bp in length
6327	6426:	gap of unknown length
6427	8873:	contig of 2447 bp in length
8874	8974:	gap of unknown length
8974	10783:	contig of 1810 bp in length
10784	10883:	gap of unknown length
10884	12763:	contig of 1880 bp in length
12764	12863:	gap of unknown length
12864	15381:	contig of 2518 bp in length
15382	15481:	gap of unknown length
15482	17365:	contig of 1884 bp in length
17366	17465:	gap of unknown length
17466	19992:	contig of 2527 bp in length
19993	20092:	gap of unknown length
20093	23512:	contig of 3420 bp in length
23513	23612:	gap of unknown length
23613	25699:	contig of 2087 bp in length
25700	25799:	gap of unknown length
25800	28333:	contig of 2534 bp in length
28334	28433:	gap of unknown length
28434	32027:	contig of 3594 bp in length
32028	32127:	gap of unknown length
32128	36115:	contig of 3988 bp in length
36116	36215:	gap of unknown length
36216	39882:	contig of 3667 bp in length

```
misc_feature 1. .1347 /note="assembly_name:Contig26"  
misc_feature 1348. .2752 /note="assembly_name:Contig31"  
misc_feature 2883. .4773 /note="assembly_name:Contig34"  
misc_feature 4874. .636 /note="assembly_name:Contig35"  
misc_feature 6427. .8873 /note="assembly_name:Contig36"  
misc_feature 8974. .10783 /note="assembly_name:Contig37"  
misc_feature 10884. .12763 /note="assembly_name:Contig38"  
misc_feature 12864. .15381 /note="assembly_name:Contig39"  
misc_feature 15482. .17365 /note="assembly_name:Contig40"  
misc_feature 17466. .19992 /note="assembly_name:Contig41"  
misc_feature 20093. .23512 /note="assembly_name:Contig42"  
misc_feature 23613. .25699 /note="assembly_name:Contig43"  
misc_feature 25800. .28333 /note="assembly_name:Contig44"  
misc_feature 28434. .32027 /note="assembly_name:Contig45"  
misc_feature 32128. .36115 /note="assembly_name:Contig46"  
misc_feature 36216. .39882 /note="assembly_name:Contig47"
```

Query Match	Similarity	29.0%	Score 220	DB 2	Length 177744
Best Local	Similarity	100.0%	Frd. No. 4,4e-99		
Match	220	Conservative	0	Mismatches	0
				Indels	0
				Gaps	0
QY	1	TACTGGAAGGTGGCGCTGCCCTCTCTGTGCTGTGATCCATGCAAGCTCCCACTGGCCCTGTGT	60		
Db	46065	TACTGGAAGGTGGCGCTGCCCTCTCTGTGCTGTGATCCATGCAAGCTCCCACTGGCCCTGTGT	46124		
QY	61	CTCGCTGCTGCTGCTGCTGTACACACAGCCTTCCGTCGTAGTGAAGGCGCAGGGGTGGACGGCG	120		
Db	46125	CTCGCTGCTGCTGCTGCTGTACACACAGCCTTCCGTCGTAGTGAAGGCGCAGGGGTGGACGGCG	46184		
QY	121	TTCAAGATGATGTCACACGGAATCATCCCCGAGCTTCGGAAGTATCCCGAGGCTCCACCG	180		
Db	46185	TTCAAGATGATGTCACACGGAATCATCCCCGAGCTTCGGAAGTATCCCGAGGCTCCACCG	46244		
QY	181	GAGCTGGAAGACACACAGACCATGAACCGGCGGAGAAAC	220		
Db	46245	GAGCTGGAAGACACACAGACCATGAACCGGCGGAGAAAC	46284		
RESULT 11					
LOCUS	AF326742	642 bp	mRNA	linear	PRI 28-FEB-2001
DEFINITION	Cercopithecus aethiops sclerostlin mRNA, complete cds.				
ACCESSION	AF326742				
VERSION	AF326742.1	GI:13161028			
KEYWORDS					
SOURCE	Cercopithecus aethiops.				
ORGANISM	Cercopithecus aethiops.				
	Bukariyota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				

FEATURES

Location/Qualifiers

1. .33790

source

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="17"

/map="17q21-q22"

<1. .1612

/gene="SOST"

/gene="SOST"

<1. .1612

/gene="SOST"

3' UTR

1. .1612

/gene="SOST"

12298. .12676

/standard_name="D17S951"

36239. .87958

/note="Van Buchem causative deletion"

79495. .79823

/standard_name="D17S1789"

93762. .>93790

/gene="MEOX1"

93762. .>93790

/gene="MEOX1"

93762. .93790

/gene="MEOX1"

5' UTR

/gene="MEOX1"

BASE COUNT 24072 a 23664 c 22642 g 23409 t 3 others

ORIGIN

Query Match 10.7%; Score 81; DB 9; Length 93790;

Best Local Similarity 100.0%; Pred. No. 2.2e-29; Indels 0; Gaps 0;

Matches 81; Conservative 0; Mismatches 0;

OY 679 AGCCGCGCCGCGCCCTCCACCGCGCGCGCCGCGCCCTTGAACCCGCGCCCAATT 738

DB 1 AGCCGCGCCGCGCCCTCCACCGCGCGCGCCGCGCCCTTGAACCCGCGCCCAATT 60

OY 739 TCTGTCTCTGCGCGCTGTT 759

DB 61 TCTGTCTCTGCGCGCTGTT 81

RESULT 14

AX323455 636 bp DNA linear PAT 07-JAN-2002

LOCUS AX323455 3 from Patent WO0192308.

DEFINITION AX323455

ACCESSION AX323455

VERSION AX323455.1 GI:18094217

KEYWORDS

SOURCE house mouse.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 Paszty,C.J. and Gao,Y.

Cystine-knot polypeptides: cloaked-2 molecules and uses thereof

Patent: WO 0192308-A 3 06-DEC-2001;

JOURNAL Amgen, Inc. (US)

FEATURES

Location/Qualifiers

1. .636

/organism="Mus musculus"

/db_xref="taxon:10090"

BASE COUNT 114 a 224 c 207 g 91 t

ORIGIN

Query Match 9.1%; Score 69; DB 6; Length 636;

Best Local Similarity 100.0%; Pred. No. 2.3e-23; Indels 0; Gaps 0;

Matches 69; Conservative 0; Mismatches 0;

OY 517 CTGTGGCTGTGAGTGAAGGCTCAACCGCTTCAACAACGAGTGAAGTCAAG 576

DB 475 CTGTGGCTGTGAGTGAAGGCTTCAACAACGAGTGAAGTCAAG 534

OY 577 GACTTCGGG 585

DB 535 GACTTCGGG 543

RESULT 15

AF326740 638 bp mRNA linear ROD 28-FEB-2001

LOCUS AF326740

DEFINITION Mus musculus sclerostin mRNA, complete cds.

ACCESSION AF326740

VERSION AF326740.1 GI:13161022

KEYWORDS

SOURCE

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 638)

Brunkow,M.E., Gardner,J.C., Van Ness,J., Paeppe,B.W., Kovacevich,B.R., Prohl,S., Skonier,J.E., Zhao,L., Sabo,P.J., Fu,Y.H., Alisch,R.S., Gillett,L., Colbert,T., Tacconi,P., Galas,D., Hamersma,H., Beighton,P. and Mulligan,J.T.

Bone dysplasia sclerosteosis results from loss of the SOST gene product, a novel cystine knot-containing protein

Am. J. Hum. Genet. 68 (3), 577-589 (2001)

JOURNAL

MEDLINE

PUBMED

21090529

REFERENCE

AUTHORS

2 (bases 1 to 638)

Brunkow,M.E., Gardner,J.C., Van Ness,J., Paeppe,B.W., Kovacevich,B.R., Prohl,S., Skonier,J.E., Zhao,L., Sabo,P.J., Fu,Y.H., Alisch,R.S., Gillett,L., Colbert,T., Tacconi,P., Galas,D., Hamersma,H., Beighton,P. and Mulligan,J.T.

Direct Submission

Submitted (07-DEC-2000) Genomics, Celltech Chiroscience Inc., 1631 220th St. SE, Bothell, WA 98021, USA

FEATURES

Location/Qualifiers

1. .638

/organism="Mus musculus"

/strain="129Sv/J"

/db_xref="taxon:10090"

/chromosome="11"

1. .636

/codon_start=1

/product="sclerostin"

/protein_id="AAK13455.1"

/db_xref="GI:13161023"

/translation="MQPSLAPPCILVHAFAVAGVGOMAFRNDATVIVPLGEYP EPPENNQTMRANIGRPPHNPYDAKGVSEYSCRLHYTRFLTBGCSARVTELV CSQCGPRLPLNAGIKVXWPRNGDPFCTIPRYAQVQLCBGGAAPRRKXRLV ASCKCKRLTFHNSSEIKDPGPEPTARPKGRKRPGARAKANQALENAY"

BASE COUNT 114 a 224 c 209 g 91 t

ORIGIN

Query Match 9.1%; Score 69; DB 10; Length 638;

Best Local Similarity 100.0%; Pred. No. 2.3e-23; Indels 0; Gaps 0;

Matches 69; Conservative 0; Mismatches 0;

OY 517 CTGTGGCTGTGAGTGAAGGCTCAACCGCTTCAACAACGAGTGAAGTCAAG 576

DB 475 CTGTGGCTGTGAGTGAAGGCTTCAACAACGAGTGAAGTCAAG 534

OY 577 GACTTCGGG 585

DB 535 GACTTCGGG 543

Search completed: March 29, 2003, 09:56:49

Job time : 2587.25 secs

PF 29-MAY-2001; 2001WO-US17478.
 XX 01-JUN-2000; 2000US-20850P.
 PR 04-AUG-2000; 2000US-223542P.
 XX (AMGE-) AMGEN INC.
 PA
 PI Paszty CJ, Gao Y;
 XX
 XX WPI: 2002-114325/15.
 DR P-PSDB; ABB07207, ABB07209.
 XX
 XX New human and mouse cysteine-knot polypeptide designated as Cloaked-2,
 PT for treating or preventing kidney, heart (e.g. myocardial infarction)
 PT or liver (e.g. hepatitis) diseases
 XX
 XX Claim 1; Fig 1; 170pp; English.

XX The invention relates to polypeptides comprising a cysteine knot motif
 CC and designated as Cloaked-2, derived from human and mouse. The Cloaked-2
 CC polypeptides can be expressed by standard recombinant methodology. The
 CC Cloaked-2 polynucleotides are useful in gene therapy and antisense
 CC therapy. The Cloaked-2 polypeptides and polynucleotides are useful for
 CC treating, preventing, ameliorating or detecting diseases and disorders of
 CC the kidney (e.g. anemia, hypertension or low blood pressure), heart (e.g.
 CC cardiac hypertrophy, congestive heart failure, myocardial infarction,
 CC arrhythmias, atherosclerosis, hypertension or low blood pressure),
 CC skeletal muscle (e.g. muscular dystrophy or cachexia), placenta (e.g.
 CC congenital abnormalities or miscarriage), liver (e.g. hepatitis or
 CC cirrhosis), pancreas (e.g. diabetes or pancreatitis), thyroid (e.g.
 CC Grave's disease or myxedema) or adrenal cortex (e.g. Cushing's disease
 CC or Addison's disease), homeostasis or metabolic diseases (e.g. obesity,
 CC cancer or myopathies), infections, or autoimmune diseases. Selective
 CC binding agents may be used to modulate the biological activities of a
 CC Cloaked-2 polypeptides or to detect Cloaked-2 polypeptide levels in a
 CC sample. Transgenic non-human animals are useful for drug candidate
 CC screening. The present sequence represents a cDNA encoding the human
 CC Cloaked-2 polypeptide.

XX Sequence 759 BP; 125 A; 282 C; 244 G; 108 T; 0 other;

XX Query Match 100.0%; Score 759; DB 24; Length 759;
 XX Best Local Similarity 100.0%; Pred. No. 0;
 XX Matches 759; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TACTGGAAGGTGGCGTCCCTCTCTGCTGTACATGCACTCCCACTGGCCCTGTGT 60
 DB 1 TACTGGAAGGTGGCGTCCCTCTCTGCTGTACATGCACTCCCACTGGCCCTGTGT 60
 QY 61 CTGCTGCGCCGTGTGTACACAGAGCCCTCCGTGTGTGTGTGTGTGTGTGTGTGT 120
 DB 61 CTGCTGCGCCGTGTGTACACAGAGCCCTCCGTGTGTGTGTGTGTGTGTGTGTGT 120
 QY 121 TTCAAGATGATGTCAGAGAAATCATCCCGAGCTCGAGAGTACCCCGAGCTCCACCG 180
 DB 121 TTCAAGATGATGTCAGAGAAATCATCCCGAGCTCGAGAGTACCCCGAGCTCCACCG 180
 QY 181 GAGCTGAGAACAAACAGACCATGAAACCGGCGGAGAAACGAGGCGCGCTCCCAACAC 240
 DB 181 GAGCTGAGAACAAACAGACCATGAAACCGGCGGAGAAACGAGGCGCGCTCCCAACAC 240
 QY 241 CCCCTTGAACAAAGAGTGTCCGAGTACAGCTGCGCGAGCTGCACTTCAACCCGCTAC 300
 DB 241 CCCCTTGAACAAAGAGTGTCCGAGTACAGCTGCGCGAGCTGCACTTCAACCCGCTAC 300
 QY 301 GTGACCGATGGGCGGTGCGAGCGCCAAAGCCGATCAACGAGTGTGTGTGTGTGTGTGT 360
 DB 301 GTGACCGATGGGCGGTGCGAGCGCCAAAGCCGATCAACGAGTGTGTGTGTGTGTGTGT 360
 QY 361 TGGCGCCGCGCGCTGT 420
 DB 361 TGGCGCCGCGCGCTGT 420

QY 421 GGGCCGACCTTCCGCTGTACATCCCGACCGCTACCGCGCGCAGCGGTGAGCTGTGTGT 480
 DB 421 GGGCCGACCTTCCGCTGTACATCCCGACCGCTACCGCGCGCAGCGGTGAGCTGTGTGT 480
 QY 481 CCCGCTGT 540
 DB 481 CCCGCTGT 540
 QY 541 CGCTTCAACCGCTTCCCAACACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 600
 DB 541 CGCTTCAACCGCTTCCCAACACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 600
 QY 601 CCGCAGAAAGGCGGAGAGCCCGCGCCCGCGCCCGAGCGCAACCAACAGCGCGAG 660
 DB 601 CCGCAGAAAGGCGGAGAGCCCGCGCCCGCGCCCGAGCGCAACCAACAGCGCGAG 660
 QY 661 CTGAGAAACGCTTCTAGAGCCCGCGCCCGCTCCCAACCGCGCGCGCGCGCGCT 720
 DB 661 CTGAGAAACGCTTCTAGAGCCCGCGCCCGCTCCCAACCGCGCGCGCGCGCGCT 720
 QY 721 GAACCCGCGCCCAATTTCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 759
 DB 721 GAACCCGCGCCCAATTTCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 759

RESULT 2

AAA29055 standard; cDNA; 2301 BP.

AAA29055;

12-SEP-2000 (first entry)

Human TGF-beta binding protein (BBER) cDNA.

osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
 BBER; chromosome 17q12-21; gene therapy; antisense therapy; fracture;
 bone mineralization; ss.

Homo sapiens.

Key Location/Qualifiers
 CDS 48..689
 FT /*tag= a
 FT /product= TGF-beta_binding_protein

W0200032773-A1.

08-JUN-2000.

24-NOV-1999; 99WO-US27990.

27-NOV-1998; 98US-0110283.

(DAFW-) DARWIN DISCOVERY LTD.

Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepker BW;
 Van Ness J, Winkler DG;

WPI: 2000-412321/35.

P-PSDB; AA196429.

Nucleic acids (1) encoding a transforming growth factor beta binding
 protein, useful for identifying agents for treating osteopenia,
 osteoporosis and fractures

Claim 1; Page 114-115; 162pp; English.

This cDNA encodes a human transforming growth factor-beta (TGF-beta)
 binding protein designated BBER. The hBBER gene has been localized
 to the chromosome 17q12-21. The cDNA and protein may be used for
 prevention, treatment and diagnosis of diseases associated with
 inappropriate BBER expression. For example, they may be used to treat

QY	3	CTGGAAAGGTGCGAGCCCTCTCTGGGCTGGTACCATGACATCCACATCGACCTGTGTCT	62
Db	28	CTGGAAAGGTGCGGTGTCCTCTCTGGCTGGTACCATGACATCCACATCGACCTGTGTCT	87
QY	63	CGTCTGCTGTGTGTACACAGACCTTCCGTGTAGTGAAGGACGAGGGTGGAGGCGTT	122
Db	88	CGTCTGCTGTGTGTGTACACAGACCTTCCGTGTAGTGAAGGACGAGGGTGGAGGCGTT	147
QY	123	CAAGAAATGATGCGACCGGAATCATCCCGGACGTGGAGAGTACCCGAGGCTCCACCGGA	182
Db	148	CAAGAAATGATGCGACCGGAATCATCCCGGACGTGGAGAGTACCCGAGGCTCCACCGGA	207
QY	183	GCTGGAGAAACAACAAGACCATGTAACCGGGCGAGAAACGAGAGGGCGGCTCCACACACC	242
Db	208	GCTGGAGAAACAACAAGACCATGTAACCGGGCGAGAAACGAGAGGGCGGCTCCACACACC	267
QY	243	CTTTGAGACCAAAAGACGTGTCCGAGTACAGCTGCCCGGAGCTGCATTACCCGCTACGT	302
Db	268	CTTTGAGACCAAAAGACGTGTCCGAGTACAGCTGCCCGGAGCTGCATTACCCGCTACGT	327
QY	303	GACCGATGAGGCGCTGGCGGACGCGCAACCGGCTACACGAGTGATGAGCTCCGGCCAGTG	362
Db	328	GACCGATGAGGCGCTGGCGGACGCGCAACCGGCTACACGAGTGATGAGCTCCGGCCAGTG	387
QY	363	CGGACCCGCGCGCTGTGTGCCCAACGCCATCGGCGCGGCAAGTGATGTGCGACTAGTGG	422
Db	388	CGGACCCGCGCGCTGTGTGCCCAACGCCATCGGCGCGGCAAGTGATGTGCGACTAGTGG	447
QY	423	GCCCGCATTTCCGCTGTGCATCCCGGACCGCTACCGGCGGAGGCGGTGACGTGTGTCC	482
Db	448	GCCCGCATTTCCGCTGTGCATCCCGGACCGCTACCGGCGGAGGCGGTGACGTGTGTCC	507
QY	483	CGGTGTGAGGAGCGCGCGCGCGCAAGTGGCGCTGTGTGACTGTGTGAAGTGCAGACG	542
Db	508	CGGTGTGAGGAGCGCGCGCGCGCAAGTGGCGCTGTGTGACTGTGTGAAGTGCAGACG	567
QY	543	CCTCAACCGGCTTCCACAACAACAGTCGAGCTCAAGGACTTTCGGGACCGAGGCGGTGGGCC	602
Db	568	CCTCAACCGGCTTCCACAACAACAGTCGAGCTCAAGGACTTTCGGGACCGAGGCGGTGGGCC	627
QY	603	GCAGAAAGGCGCGGAAGCGCGGCGCCCGCGCTCGAGAGCGCCAAAGCCAAAGGCGGAGCT	666
Db	628	GCAGAAAGGCGCGGAAGCGCGGCGCCCGCGCTCGAGAGCGCCAAAGGCGGCGGAGCT	687
QY	663	GGAGAACGCTACTGTAGAGCGCGCGCGCGCTCCCGACCGGCGGAGCGCCCGGCGCTGA	722
Db	688	GGAGAACGCTACTGTAGAGCGCGCGCGCGCTCCCGACCGGCGGAGCGCCCGGCGCTGA	747
QY	723	ACCCGCGCCCAACATTTCTGTCTCTGTGCGCGTGTATT	759
Db	748	ACCCGCGCCCAACATTTCTGTCTCTGTGCGCGTGTATT	784
RESULT 4			
ABK69992			
ID	ABK69992 standard; DNA; 2329 BP.		
XX	ABK69992;		
AC	15-JUL-2002 (first entry)		
DT	cDNA encoding human Pro peptide #32.		
XX	Human; ss; gene; PRO; secreted protein; transmembrane protein;		
DE	genetic disorder; tumour; cancer.		
XX	Homo sapiens.		
OS	WO200224888-A2.		
PN	28-MAR-2002.		
XX			
PD			

PF	29-AUG-2001;	2001WO-US27099.
XX		
PR	01-SEP-2000;	2000US-229896P.
PR	05-SEP-2000;	2000US-230621P.
PR	22-SEP-2000;	2000US-235147P.
PR	10-NOV-2000;	2000WO-US30873.
PR	12-JAN-2001;	2001US-261878P.
PR	16-JAN-2001;	2001US-261910P.
PR	16-JAN-2001;	2001US-261939P.
PR	16-JAN-2001;	2001US-262150P.
PR	25-JAN-2001;	2001US-264395P.
PR	02-FEB-2001;	2001US-266421P.
PR	09-FEB-2001;	2001US-267623P.
PR	28-FEB-2001;	2001WO-US06520.
PR	09-MAR-2001;	2001US-274399P.
PR	03-APR-2001;	2001US-280982P.
PR	04-APR-2001;	2001US-282129P.
PR	04-APR-2001;	2001US-282199P.
PR	09-MAY-2001;	2001US-290589P.
PR	25-MAY-2001;	2001WO-US17092.
PR	01-JUN-2001;	2001WO-US17800.
PR	20-JUN-2001;	2001WO-US19692.
PR	29-JUN-2001;	2001WO-US21066.
PR	09-JUL-2001;	2001WO-US21735.

Baker KP, Eaton DL, Filvaroff E, Goddard A, Grimaldi JC, Gurney AL, Smith V, Stephan J, Watanabe CK, Wood WI, Zhang Z, Fong S;
MPI: 2002-362426/39.
P-PSDB; ABG34061.

New PRO polypeptides and polynucleotides encoding the polypeptides, useful in gene therapy, chromosome identification, tissue typing, or for genetic analysis of individuals with genetic disorders -

This invention relates to the cDNA and protein sequences of novel secreted and transmembrane polypeptides PRO polypeptides. The invention also comprises a method for producing the proteins of the invention by recombinant means and antibodies specific for the protein of the invention. The antibody may be used for detecting the PRO proteins of the invention and may be used to modify their activity. polynucleotides may be used as hybridisation probes for a cDNA library to isolate the full-length PRO cDNA or to isolate other cDNAs, to construct hybridisation probes for mapping the gene which encodes that PRO and for genetic analysis of individuals with genetic disorders, in assays to identify other proteins or molecules involved in binding reaction, to generate transgenic animals or knock-out animals which in turn are useful in the development and screening of therapeutically useful reagents, for chromosome identification, and as molecular weight PRO polypeptides are useful in gene therapy, and as molecular weight markers for protein electrophoresis purposes. The sequences may also be used to detect overexpression on PRO polypeptides in cancerous tumours and for screening for differentially expressed genes using microarray technology. The present sequence represents a cDNA encoding a human PRO protein of the invention.

SQ Sequence 2329 BP; 587 A; 645 C; 612 G; 485 T; 0 other;

Query Match	Score	DB	Length
99.7%	757	24	2329

```
Matches 757; Conservative 0; Mismatches 0; Indels 0; Gaps 0
```

QY 3 CTGGAAGGTGGGCTGCCCTCTCTGTGGCTGTACCATGATAGCTCCCACTGGACCGCTGTCT 62

Db 28 CTGGAAGGTGGGCTGCCCTCTCTGTGGCTGTACCATGATAGCTCCCACTGGACCGCTGTCT 87

QY 63 CGTCTGCTGTGTATCACACAGCTTTCGCTGTATGTGAGGGCCAGGGGTGGACAGGGTT 122

Db 361 AACCCATCGCCCGGGAAGTGTGGGACCTAGTGGGCCGACCTTCCGCTGCATCCCC 420
 Qy 445 GACCGCTACCGCGCGGACCGGTGCTGTGTGTCCCGTGTGAGGGCGCGCGCG 504
 Db 421 GACCGCTACCGCGCGGACCGGTGCTGTGTGTCCCGTGTGAGGGCGCGCGCG 480
 Qy 505 CGGAGGCGCGCGGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 564
 Db 481 CGGAGGCGCGCGGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 540
 Qy 565 TCGAGCTCAAGAGCTTGGGAGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 624
 Db 541 TCGAGCTCAAGAGCTTGGGAGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
 Qy 625 CCG 684
 Db 601 CCG 660
 Qy 685 CCG 744
 Db 661 CCG 720
 Qy 745 CTCTGCGCGTGTGT 759
 Db 721 CTCTGCGCGTGTGT 735

RESULT 6
 AAA29061
 ID AAA29061 standard; cDNA; 2301 BP.

XX AAA29061;
 XX 12-SEP-2000 (first entry)
 XX
 XX Mutant human TGF-beta binding protein (BBER) cDNA.
 XX
 XX osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
 XX BBER; chromosome 17q12-21; gene therapy; antisense therapy; fracture;
 XX bone mineralization; mutant; sclerosteosis; ss.
 XX Homo sapiens.
 OS

XX Key Location/Qualifiers
 XX FH 48.119
 XX CDS /*tag= a
 XX FT /note= "mutant BBER created by sclerosteosis
 XX FT /note= "nonsense mutation"

XX MO200032773-A1.
 XX 08-JUN-2000.
 XX PD
 XX 24-NOV-1999; 99MO-US27990.
 XX PR 27-NOV-1998; 98US-0110283.
 XX PA (DARW-) DARWIN DISCOVERY LTD.
 XX PI Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepier BW;
 XX PI Van Ness J, Winkler DG;
 XX DR WPI; 2000-412321/35.
 XX DR P-PSDB; AAY96435.
 XX Nucleic acids (1) encoding a transforming growth factor beta binding
 XX PT protein, useful for identifying agents for treating osteopenia,
 XX PT osteoporosis and fractures
 XX PS Claim 27; Page 117-118; 162pp; English.
 XX CC This cDNA encodes a mutant human transforming growth factor-beta

CC (TGF-beta) binding protein (BBER) produced as a result of a nonsense
 CC mutation in the BBER coding sequence (C to T mutation at position 117).
 CC This mutation has been linked to sclerosteosis in humans by linkage
 CC analysis of an Afrikaaner population in South Africa. The hBBER gene has
 CC been localized to the chromosome 17q12-21. The cDNA and protein may be
 CC used for prevention, treatment and diagnosis of diseases associated with
 CC inappropriate BBER expression. For example, they may be used to treat
 CC disorders associated with decreased TGF-beta BP expression. The cDNA or
 CC vectors may be administered to treat diseases by rectifying mutations or
 CC deletions in a patient's genome that affect the activity of BBER by
 CC expressing inactive proteins or to supplement the patient's own production
 CC of BBER polypeptides. The nucleic acids may be used for recombinant
 CC production of BBER, gene therapy, antisense therapy, as probes for
 CC diagnostic assays and for functional studies. BBER may be used to raise
 CC antibodies and for identification of BBER modulators. BBER antagonists
 CC may be used to increase bone mineral content for the treatment of
 CC disorders such as osteopenia, osteoporosis, fractures and other
 CC disorders associated with low mineral content.

SQ Sequence 2301 BP; 568 A; 634 C; 614 G; 485 T; 0 other;

Query Match 93.3%; Score 708; DB 21; Length 2301;

Best Local Similarity 99.9%; Pred. No. 0; Mismatches 1; Indels 0; Gaps 0;

Db 1 TACTGGAAGGTGGCGTGGCTCTCTGCTGTGATCCATGACCTCCACTGGCCCTGTGT 60
 Qy 12 TACTGGAAGGTGGCGTGGCTCTCTGCTGTGATCCATGACCTCCACTGGCCCTGTGT 71
 Db 61 CTGCTGTGCTGTGCTGTGATCAGACAGCTTCCGCTGTGTGTGTGTGTGTGTGTGTGT 120
 Qy 72 CTGCTGTGCTGTGCTGTGATCAGACAGCTTCCGCTGTGTGTGTGTGTGTGTGTGTGT 131
 Db 121 TTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 180
 Qy 132 TTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 191
 Db 181 GAGCTGAGAGCAACAAGACCATGATGATGATGATGATGATGATGATGATGATGATG 240
 Qy 192 GAGCTGAGAGCAACAAGACCATGATGATGATGATGATGATGATGATGATGATGATG 251
 Db 241 CCCCTTGAAGCAACAAGACCATGATGATGATGATGATGATGATGATGATGATGATG 300
 Qy 252 CCCCTTGAAGCAACAAGACCATGATGATGATGATGATGATGATGATGATGATGATG 311
 Db 301 GTGACCGATGAGCG 360
 Qy 312 GTGACCGATGAGCG 371
 Db 372 TGCGGCG 431
 Qy 361 TGCGGCG 420
 Db 421 GGGCG 480
 Qy 432 GGGCG 491
 Db 481 CCGGCTGTGAGAGCG 540
 Qy 492 CCGGCTGTGAGAGCG 551
 Db 541 CGCCTGACCG 600
 Qy 552 CGCCTGACCG 611
 Db 601 CCGGAGAGAGCG 660
 Qy 612 CCGGAGAGAGCG 671
 Db 661 CTGAGAGAGCGCTTACTAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 720
 Qy 672 CTGAGAGAGCGCTTACTAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 731

FT /product= tgf-beta_binding_protein
 XX MO200032773-A1.
 PN 08-JUN-2000.
 PD 24-NOV-1999; 99WO-US27990.
 PF 27-NOV-1998; 98US-0110283.
 XX (DARW-) DARWIN DISCOVERY LTD.
 PA Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepfer BW;
 PI Van Ness J, Winkler DG;
 XX MPI; 2000-412321/35.
 DR P-PSDB; AAY96430.
 XX
 PT Nucleic acids (1) encoding a transforming growth factor beta binding
 PT protein, useful for identifying agents for treating osteopenia,
 PT osteoporosis and fractures
 PS Claim 1; Page 118-119; 162pp; English.
 CC This cDNA encodes a variant human transforming growth factor-beta
 CC (TGF-beta) binding protein designated BEER VI01. The encoded protein
 CC comprises a substitution of isoleucine for the wild-type valine at
 CC residue 10. The cDNA and protein may be used for prevention, treatment
 CC and diagnosis of diseases associated with inappropriate BEER expression.
 CC For example, they may be used to treat disorders associated with
 CC decreased TGF-beta BP expression. The cDNA or vectors may be administered
 CC to treat diseases by rectifying mutations or deletions in a patient's
 CC genome that affect the activity of BEER by expressing inactive proteins
 CC or to supplement the patient's own production of BEER polypeptides. The
 CC nucleic acids may be used for recombinant production of BEER, gene
 CC therapy, antisense therapy, as probes for diagnostic assays and for
 CC functional studies. BEER may be used to raise antibodies and for
 CC identification of BEER modulators. BEER antagonists may be used to
 CC increase bone mineral content for the treatment of disorders such as
 CC osteopenia, osteoporosis, fractures and other disorders associated with
 CC low mineral content.
 CC
 CC Sequence 2301 BP; 569 A; 634 C; 614 G; 484 T; 0 other;
 SQ
 Query Match 86.6%; Score 657; DB 21; Length 2301;
 Best Local Similarity 99.7%; Pred. No. 1.6e-295;
 Matches 757; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 TACTGGAAGTGGCGTCCCTCTCTGAGTGTACATGACAGCTCCCACTGGCCCTGTGT 60
 DB 12 TACTGGAAGTGGCGTCCCTCTCTGAGTGTACATGACAGCTCCCACTGGCCCTGTGT 71
 QY 61 CTGATCTGCTGCTGTGTACACACAGCTTCCGTGTAGTGAAGGGCCAGGGGTGGCAGGGC 120
 DB 72 CTGATCTGCTGCTGTGTACACACAGCTTCCGTGTAGTGAAGGGCCAGGGGTGGCAGGGC 131
 QY 121 TTCAAGAAATGATGCGACGGAATCATCCCGAGCTCGGAGAGTACCCCGAGCTCCACCG 180
 DB 132 TTCAAGAAATGATGCGACGGAATCATCCCGAGCTCGGAGAGTACCCCGAGCTCCACCG 191
 QY 181 GAGCTGGAACAACAAGACCATGAAACCGGCGGAGAGACGAGGGGCGCTCCCAACAC 240
 DB 192 GAGCTGGAACAACAAGACCATGAAACCGGCGGAGAGACGAGGGGCGCTCCCAACAC 251
 QY 241 CCCTTTGAGACCAAGAGCGTGTCCGAGTACAGCTGCGCGAGCTTCACTTCAACCCGCTAC 300
 DB 252 CCCTTTGAGACCAAGAGCGTGTCCGAGTACAGCTGCGCGAGCTTCACTTCAACCCGCTAC 311
 QY 301 GTGACCATGGGCGCTGCGGAGCGCCAAAGCGGTCACCCGAGCTGTGTCTCGGGCAG 360
 DB 312 GTGACCATGGGCGCTGCGGAGCGCCAAAGCGGTCACCCGAGCTGTGTCTCGGGCAG 371
 QY 361 TGGGCGCGGCGCTGTGCGCAACGACCATGGGCGCGCAAGTGTGTGCGACTTGT 420

DB 372 TGGGCGCGGCGCTGTGCGCAACGACCATGGGCGCGCAAGTGTGTGCGACTTGT 431
 QY 421 GGGGCCGACCTTCCGCTGATCTCCGACCGCTACCGGCGGAGGCGGTGAGCTGTGT 480
 DB 432 GGGGCCGACCTTCCGCTGATCTCCGACCGCTACCGGCGGAGGCGGTGAGCTGTGT 491
 QY 481 CCGGTGTGTGAGGCGCGGCGCGGCGGAGGAGTGTGTGTGTGTGTGTGTGTGTGTGT 540
 DB 492 CCGGTGTGTGAGGCGCGGCGCGGCGGAGGAGTGTGTGTGTGTGTGTGTGTGTGT 551
 QY 541 CGCTTCAACCGCTTCAACAACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 600
 DB 552 CGCTTCAACCGCTTCAACAACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 611
 QY 601 CCGCAAGAGGCGCGGAGCGCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 660
 DB 612 CCGCAAGAGGCGCGGAGCGCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 671
 QY 661 CTGAGAAAGCGCTACTAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 720
 DB 672 CTGAGAAAGCGCTACTAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 731
 QY 721 GAAACCGGCT 759
 DB 732 GAAACCGGCT 770
 RESULT 9
 ID AAA94051 standard; DNA; 642 BP.
 XX AAA94051;
 AC 15-JAN-2001 (first entry)
 DT Human DAN/Cerberus-related protein 6 (hDCR6) cDNA exons 1 and 4.
 DE Human; DAN/Cerberus-related protein 6; hDCR6; morphogenic protein;
 XX antagonist; BMP; cell growth; cell differentiation; bone formation;
 KW gene therapy; ss.
 XX
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FH 1..642
 FT CDS /*tag= a
 FT /product= "hDCR6 #2"
 FT /partial
 PN WO200055193-A2.
 XX 21-SEP-2000.
 PD 02-MAR-2000; 2000WO-US05537.
 PF 12-MAR-1999; 99US-0124118.
 PR (REG-) REGENERON PHARM INC.
 PA Economides AN;
 PI MPI; 2000-638179/61.
 DR P-PSDB; AAB26106.
 XX Novel isolated, human DNA/Cerberus related protein 6 which include
 PT natural homologue, and polypeptides comprising DCR6 domain and nucleic
 PT acids encoding the proteins which are useful as probes and primers -
 XX
 PS Claim 2; Fig 3; 40pp; English.
 CC The present sequence comprises exons 1 and 4 of the human
 CC DAN/Cerberus-related protein 6 (hDCR6) coding sequence. It was isolated

	Db	721	CYTGCGGCTGGTTT	735
RESULT 6				
AF326742			642 bp	mRNA linear PRI 28-FEB-2001
LOCUS				
DEFINITION			Cercopithecus aethiops sclerostin mRNA, complete cds.	
AF326742				
AF326742.1			GI:13161028	
KEYWORDS				
SOURCE				
ORGANISM				
			Cercopithecus aethiops.	
			Cercopithecus aethiops	
			Eukaryota; Eumetazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
			Mammalia; Euarcharia; Primates; Catarrhini; Cercopithecoidea;	
			Cercopithecinae; Cercopithecus.	
			1 (bases 1 to 642)	
REFERENCE				
AUTHORS			Brunkow,M.E., Gardner,J.C., Van Ness,J., Paeppe,B.W.,	
			Kovacevich,B.R., Proll,S., Skonier,J.E., Zhao,L., Sabo,P.J.,	
			Fu,Y.H., Alisch,R.S., Gillett,L., Colbert,T., Tacconi,P., Galae,D.,	
			Hamerema,H., Beighton,P. and Mulligan,J.T.	
TITLE			Bone dysplasia sclerosteosis results from loss of the SOST gene	
JOURNAL			product, a novel cysteine knot-containing protein	
MEDLINE			Am. J. Hum. Genet. 68 (3), 577-589 (2001)	
PUBMED			21090529	
			11179006	
REFERENCE				
AUTHORS			2 (bases 1 to 642)	
			Brunkow,M.E., Gardner,J.C., Van Ness,J., Paeppe,B.W.,	
			Kovacevich,B.R., Proll,S., Skonier,J.E., Zhao,L., Sabo,P.J.,	
			Fu,Y.H., Alisch,R.S., Gillett,L., Colbert,T., Tacconi,P., Galae,D.,	
			Hamerema,H., Beighton,P. and Mulligan,J.T.	
TITLE			Direct Submision	
JOURNAL			Submitted (07-DEC-2000) Genomics, Celltech Chiroscience Inc., 1631	
FEATURES			Location/Qualifiers	
SOURCE			1..642	
			/organism="Cercopithecus aethiops"	
			/db_xref="taxon:9534"	
CDS			1..642	
			/codon_start=1	
			/product="sclerostin"	
			/protein_id="AAK13457.1"	
			/db_xref="gi:13161028"	
			/translation="MQLPLAICLVLIQVLAHFAFVEVGQAOFKNDAIEIIPELGEYR	
			LPPPELENNKTMRNAENGSRPHHPERFKDGVSEVSCRILHTTRIVVTGPSARPYHE	
			LVSGGCGSPRALPNAIGRGKWPBSGRDFPICIDRIHAQCVOLLCEGAAPRRKRPR	
			LVASCKRKRLTRFRNOSLELKDFGDEAAPQGRKRPAPARGAKANQALENAV"	
BASE COUNT			113 a 228 c 210 g 91 t	
ORIGIN				
Query Match			80.8%; Score 613.2; DB 9; Length 642;	
Best Local Similarity			97.2%; Pred. No. 3e-84;	
Matches 624; Conservative			0; Mismatches 18; Indels 0; Gaps 0;	
Qy	37	ATGCAGCTCCCACTGGCCCTGTGTTCGTCTGCTGCTGCTGCTACACACACTTCCTGGTA	96	
Db	1	ATGCAGCTCCCACTGGCCCTGTGTTCGTCTGCTGCTGCTGCTACACACACTTCCTGGTA	60	
Qy	97	GTCGAGGGCCAGGGGTGGAGGCGTTCAAGAATGATGCCACGGAATCATCCCGAGCTC	156	
Db	61	GTCGAGGGCCAGGGGTGGAGGCGTTCAAGAATGATGCCACGGAATCATCCCGAGCTC	120	
Qy	157	GAGAGGTACCCCGAGCCCTCCACCAGGAGCTGGAGAACACAAGACATGAACCGGGCGGAG	216	
Db	121	GAGAGGTACCCCGAGCCCTCCACCAGGAGCTGGAGAACACAAGACATGAACCGGGCGGAG	180	
Qy	217	AACGAGGGGCGGCTCCCAACCACTTTTGAAGACCAAGAAGTGTCCAGATAAGCTGC	276	
Db	181	AATGAGGGGCGGCTCCCAACCACTTTTGAAGACCAAGAAGTGTCCAGATAAGCTGC	240	
Qy	277	CGCAGAGTCGACTTACCCCGCTAGTAGCCATGAGGCGGTGCGCGGACCGCACCGGCTC	336	
Db	241	CGAGAGTCGACTTACCCCGCTAGTAGCCATGAGGCGGTGCGCGGACCGCACCGGCTC	300	

QY	337	ACCGAGCTGATGTCTTCGCCGCACAGTGGCCGCCGCCGCTGTGCTGCCCAACGCATCGCC	396
Db	301	ACCAGATTGTGTGTCTCCGCGCAGTGCGGAGCCGGACA GCGCTGTGCCCCAACGCATCGCC	360
QY	397	CCGCGCAGAATGTGTGCGACACTAGTGGGCCCACACTTCCGCTGCATCCCGACCGCTACCG	456
Db	361	CCGCGCAGAATGTGTGCGCCCGAGTGGGCCCATTTCCGCTGCATCCCGACCGCTACCG	420
QY	457	GCGCAGCCGCTGCAGCTGTGTGTCTCCGATGTGAAGCCGCCGCCGCCGCACAGTGGCC	516
Db	421	GCGCAGCCGCTGCAGCTGTGTGTCTCCGATGTGAAGCCGCCGCCGCCGCACAGTGGCC	480
QY	517	CCTGTGAGCTGTGTGAAGTGAAGCAGGCGCTCACCGCTTCCACAACAGTGGAGCTCAAG	576
Db	481	CTGTGTGAGCTGTGTGAAGTGAAGCAGGCGCTCACCGCTTCCACAACAGTGGAGCTCAAG	540
QY	577	GACTTCGAGGACCGAAGCGCGCTGCGCCGCGCAGAAGGAGCCGGAACCGCGCGCCGCGCG	636
Db	541	GACTTCGAGGACCGAAGCGCGCTGCGCCGCGCAGAAGGAGCCGGAACCGCGCGCCGCGCG	600
QY	637	AGCGCCAAAGCCAACCAAGCGCGAGCTGGAGAAAGCGCTACTAG	678
Db	601	GGGCGCAAAGCCAATCAAGCGCGAGCTGGAGAAAGCGCTACTAG	642
RESULT 7			
AF326741			
LOCUS	AF326741	674 bp	mRNA linear ROD 28-FEB-2001
DEFINITION	Rattus norvegicus sclerostin mRNA, complete cds.		
ACCESSION	AF326741		
VERSION	AF326741.1	GI:13161025	
KEYWORDS	.		
SOURCE	Rattus norvegicus.		
ORGANISM	Rattus norvegicus.		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;		
	Rattus.		
REFERENCE	1 (bases 1 to 674)		
AUTHORS	Brunkow,M.E., Gardner,J.C., Van Ness,J., Pieper,B.W., Kovacevich,B.R., Proll,S., Skonier,J.E., Zhao,L., Sabo,P.J., Fu,Y.H., Altsch,R.S., Gillett,L., Colbert,T., Tacconi,P., Galas,D., Hamersma,H., Beighon,P. and Mulligan,J.T.		
TITLE	Bone dysplasia sclerosteosis results from loss of the SOST gene product, a novel cystine knot-containing protein		
JOURNAL	Am. J. Hum. Genet. 68 (3), 577-589 (2001)		
PUBMED	21090529		
REFERENCE	2 (bases 1 to 674)		
AUTHORS	Brunkow,M.E., Gardner,J.C., Van Ness,J., Pieper,B.W., Kovacevich,B.R., Proll,S., Skonier,J.E., Zhao,L., Sabo,P.J., Fu,Y.H., Altsch,R.S., Gillett,L., Colbert,T., Tacconi,P., Galas,D., Hamersma,H., Beighon,P. and Mulligan,J.T.		
TITLE	Direct Submission		
JOURNAL	Submitted (07-DEC-2000) Genomics, Celltech Chiroscience Inc., 1631		
FEATURES	220th St. SE, Bothell, WA 98021, USA		
Source	Location/Qualifiers		
	1..674	/organism="Rattus norvegicus"	
		/strain="Sprague-Dawley"	
		/db_xref="taxon:10116"	
	1.._32		
	33..674	/codon_start=1	
		/product="sclerostin"	
		/protein_id="AAK13456.1"	
		/db_xref="GI:13161026"	
		/translation="MQLSLAPCLALVHAFAVASGSGMAFKNDAETIPGLREYR EPPELENNOTNRAENGGRPHPIPTKYSEISCEELHTTRPVTPGPCSAKPVE LVSSGCGPAPLFPNALGVKVMPPNSDPDRICLPDRIADRVVLCPGGAAPRSKRVR LVASCKRKRLTRFNOSLKDVGFTIRPKGRKRPARGAKRANOAELEWAY"	
BASE COUNT	128 a	240 c	208 g 98 t
ORIGIN			

Kovacevich, B.R., Prohl, S., Skonier, J.E., Zhao, L., Sabo, P.U.,
Fu, Y.H., Altsch, R.S., Gillett, L., Colbert, T., Tacconi, P., Galas, D.,
Hamerema, H., Beighson, P. and Mulligan, J.T.
Submitted (07-DEC-2000) Genomics, Cellegch Chiroscience Inc., 1631
220th St. SE, Bothell, WA 98021, USA

TITLE JOURNAL

FEATURES

SOURCE

1. 21501
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="17"
/map="17q21"
join(11961..12180,14939..15360)
/product="sclerostin"
join(11961..12180,14939..15360)
/codon_start=1
/product="sclerostin"
/protein_id="AAK13451.1"
/db_xref="GI:13161011"
/translation="MQLPLALCLVCLLHTAFVVEGQXWQAFKDATEIIPELGEYP
EPPELENNKTMRAENGRRPHNPEITKDVSEYSCREHFTYVTDGCRSAKPYTE
LVSGGCGPARLLPNAIGRGKWRPSGDFRCIPDRYRAORVOLCPGGEAPARAVR
LVASCKKRLTRFHNOSELKDFTEAARPKRKRPRPARSAKANAELNNY"

CDs

MEWA

BASE COUNT 5546 a 5259 c 5419 g 5277 t

ORIGIN

Query Match 68.1%; Score 516.8; DB 9; Length 21501;
Best Local Similarity 96.0%; Pred. No. 6.6e-70;

Matches 530; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 208 CGGGCGGAGAACGAGAGGGCGGCTCCACACACCCCTTTAGACCAAAAGCTGTCCGAG 267
DB 14890 CAGGGGCGGAGAGGGCGGCTCCACACACCCCTTTAGACCAAAAGCTGTCCGAG 14949
QY 268 TACAGCTGCGCGAGCTGACCTTACCCCGCTACGACGAGGCGCGTGGCGGAGGCC 327
DB 14950 TACAGCTGCGCGAGCTGACCTTACCCCGCTACGACGAGGCGCGTGGCGGAGGCC 15009
QY 328 AAGCGCGTACCGAGCTGTGTGCTCCGCGCAGTGGCGGCGCGCTGTGCTGCCAAC 387
DB 15010 AAGCGCGTACCGAGCTGTGTGCTCCGCGCAGTGGCGGCGCGCTGTGCTGCCAAC 15069
QY 388 GCCATGCGCGCGCGAGTGTGTGCTCCGCGCAGTGGCGGCGCGCTGTGCTGCCAAC 447
DB 15070 GCCATGCGCGCGCGAGTGTGTGCTCCGCGCAGTGGCGGCGCGCTGTGCTGCCAAC 15129
QY 448 CGCTACCGCGCGAGCGCTGACGCTGTGTGCTCCGCGTGTGAGGCGCGCGCGCGCG 507
DB 15130 CGCTACCGCGCGAGCGCTGACGCTGTGTGCTCCGCGTGTGAGGCGCGCGCGCGCG 15189
QY 508 AAGGTGCGCTGTGTGCTGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 567
DB 15190 AAGGTGCGCTGTGTGCTGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 15249
QY 568 GAGCTCAAGACTTTCGGGACCGAGCGCGCTCGCGCGGAGAGGCGCGGAGCGCGCGCC 627
DB 15250 GAGCTCAAGACTTTCGGGACCGAGCGCGCTCGCGCGGAGAGGCGCGGAGCGCGCC 15309
QY 628 CGGCGCGGAGCGCGCAAGCGCAACGAGCGCGAGCTGAGGAACCTCTAAGAGCGCGCC 687
DB 15310 CGGCGCGGAGCGCGCAAGCGCAACGAGCGCGAGCTGAGGAACCTCTAAGAGCGCGCC 15369
QY 688 GCGCGCTTCCCGACCGGCGCGCGCGCGCGCTGGAACCGCGCGCGCACTTTCTGTCTTC 747
DB 15370 GCGCGCTTCCCGACCGGCGCGCGCGCGCGCGCTGGAACCGCGCGCGCACTTTCTGTCTTC 15429
QY 748 TCGCGGTGTTT 759
DB 15430 TCGCGGTGTTT 15441

RESULT 10
AC055813/c

LOCUS AC055813 80117 bp DNA linear PRI 22-FEB-2002
DEFINITION Homo sapiens chromosome 17, clone RP11-209M4, complete sequence.
AC055813
AC055813.9 GI:16860756
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

1 (bases 1 to 80117)
Birren, B., Linton, L., Nusbaum, C. and Lander, E.
Homo sapiens chromosome 17, clone RP11-209M4
Unpublished
2 (bases 1 to 80117)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Abramson, H., Allen, N.,
Anderson, S., Baldwin, J., Barna, N., Bastien, V., Beda, F.,
Boguslavsky, L., Bouckigalter, B., Brown, A., Burkett, G.,
Campopiano, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S.,
Collamore, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S.,
Dodge, S., Domingo, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D.,
Galagan, J., Gardyna, S., Glende, S., Goyette, M., Graham, L.,
Grand-Pierre, N., Grant, G., Hagos, B., Heatford, A., Horton, L.,
Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kam, L., Karakas, A.,
Klein, J., Lakoque, K., Lamazares, R., Landers, T., Lehoczy, J.,
Levine, R., Lieu, C., Liu, G., Locke, K., Macdonald, P., Margulis, N.,
McCarthy, M., McEwan, P., McGuirk, A., McKernan, K., McPheters, R.,
Meldrum, J., Meneus, L., Mihova, T., Miranda, C., Mleaga, V., Morrow, J.,
Murphy, D., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,
O'Neill, D., Olivier, T.M., Oliver, J., Peterson, K., Pierre, N.,
Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,
Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
Testfaye, S., Theodore, J., Tirrell, A., Travers, M., Trigglio, J.,
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,
Young, G., Zainoun, J., Zimmer, A. and Zody, M.

TITLE

JOURNAL

REFERENCE

AUTHORS

Direct Submission
Submitted (18-APR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 80117)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Bouckigalter, B.,
Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B.,
Choepel, Y., Colangelo, M., Collins, S., Collamore, A., Cook, A.,
Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Fato, S.,
Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,
Glende, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Heatford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,
Jones, C., Kamat, A., Karakas, A., Kells, C., Lakoque, K.,
Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G.,
Maclean, C., Macdonald, P., Major, J., Margulis, N., Matthews, C.,
McCarthy, M., McEwan, P., McKernan, K., McPheters, R., Meldrum, J.,
Meneus, L., Mihova, T., Mleaga, V., Murphy, T., Naylor, J., Nguyen, C.,
Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D.,
Olivier, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
Raymond, C., Retta, R., Riedack, M., Riley, R., Rise, C., Rogov, P.,
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupack, R.,
Seaman, S., Severy, P., Sougnuez, C., Spencer, B., Stange-Thomann, N.,
Stojanovic, N., Straus, N., Subramanian, A., Talamas, J., Testfaye, S.,
Theodore, J., Topham, K., Travers, M., Travis, N., Trigglio, J.,
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,
Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE

JOURNAL

REFERENCE

AUTHORS

Direct Submission
Submitted (01-AUG-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 80117)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Bouckigalter, B.,
Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B.,
Choepel, Y., Colangelo, M., Collins, S., Collamore, A., Cook, A.,
Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Fato, S.,
Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,
Glende, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,

Kamat, A., Kartas, A., Kelle, C., Lacroque, K., Lamazares, R., Landers, T., Lenoczky, J., Levine, R., Liu, G., Maclean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Menes, L., Mihova, T., Mlenka, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retter, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vasilev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zaitoun, J., Zembek, L., Zimmer, A. and Zody, M.

Center: Whitehead Institute/ MIT Center for Genome Research
Genome Center
Web site: <http://www-seg.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
Project Information
Center project name: L7949
Center clone name: 209_M_4

Direct Submission
Submitted (22-FEB-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Feb 22, 2002 this sequence version replaced gi:15055348.
All repeats were identified using RepeatMasker:
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>
Smit, A.F.A. & Green, P. (1996-1997)

Only the last 80.1 kilobases of this clone are being submitted.
The remainder overlaps accession number AC004149 [WICR project L302].

FEATURES

source

Location/Qualifiers
1. 80117
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="17"
/map="17"
/clone="RP11-209M4"
/clone_lib="RP11-11 Human Male BAC"
/rpt_family="AluSg"
complement(146.434)
repeat_region
complement(438.655)
/rpt_family="MIR"
complement(974.1129)
repeat_region
complement(1270.1438)
/rpt_family="AluJo"
1439.1465
repeat_region
/rpt_family="(TTTAA)n"
complement(146.1584)
repeat_region
/rpt_family="AluJo"
complement(1785.1933)
repeat_region
/rpt_family="MERSB"
2339.2487
repeat_region
/rpt_family="MIR"
complement(2492.2787)
repeat_region
/rpt_family="AluSg"
complement(2796.2902)
repeat_region
/rpt_family="MIR"
complement(2903.3165)
repeat_region
/rpt_family="AluSg"
complement(3166.3193)
repeat_region
/rpt_family="MIR"
3572.3644
repeat_region
/rpt_family="MIR"
4524.4583
repeat_region
complement(4755.4962)
/rpt_family="MIR"
6397.6738
repeat_region

repeat_region /rpt_family="L2"
complement(6766.6875)
/rpt_family="MLT1H"
repeat_region 7202.7270
/rpt_family="L2"
repeat_region 7739.7825
/rpt_family="L2"
complement(7826.8122)
/rpt_family="AluSc"
repeat_region 8123.8268
/rpt_family="L1MC4"
repeat_region 8269.8568
/rpt_family="AluJo"
repeat_region 8569.8631
/rpt_family="L1MC4"
complement(9151.9305)
/rpt_family="MIR"
repeat_region 9365.9390
/rpt_family="AT_rich"
complement(9405.9708)
/rpt_family="AluY"
complement(9709.9969)
/rpt_family="AluJo"
repeat_region 10197.10458
/rpt_family="MLT1K"
repeat_region 10758.10895
/rpt_family="MIR"
repeat_region 12676.12773
/rpt_family="MIR"
repeat_region complement(12784.12997)
/rpt_family="MIR"
repeat_region complement(13220.13513)
/rpt_family="AluJo"
repeat_region 13627.13787
/rpt_family="L2"
complement(13821.13953)
/rpt_family="MIR"
repeat_region 14746.14899
/rpt_family="MIR"
repeat_region complement(14947.15144)
/rpt_family="MIR"
repeat_region 15520.15668
/rpt_family="AluJo"
repeat_region 15669.15960
/rpt_family="AluSx"
repeat_region 15961.16128
/rpt_family="AluJo"
repeat_region 16208.16400
/rpt_family="MIR"
repeat_region 16454.16586
/rpt_family="L2"
repeat_region 16707.16783
/rpt_family="MIR"
repeat_region 16852.16909

Query Match 68.1%; Score 516.8; DB 9; Length 80117;
Best Local Similarity 96.0%; Pred. No. 5e-70;
Matches 530; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 208 CCGGCGGAGAGAGAGGCGGCGCTCCCAACACCCCTTTAGACCAAGAGTGTCCGAG 267
DB 61263 CAGGGGTGGCCAGGGGCGGCGCACCTTCACGCGGCTCTCCACAGAGTGTCCGAG 63204
QY 268 TACAGTGCAGGAGCTGACCTTACCCGCTACGTAACCGATGGGCGCTGCGGAGCGCC 327
DB 61203 TACAGTGCAGGAGCTGACCTTACCCGCTACGTAACCGATGGGCGCTGCGGAGCGCC 63144
QY 328 AAGCCGTCACCGAGCTGCTGCTCGGCGAGTGGCGGCGCGCGCTCTGCTGCGCAAC 387
DB 61143 AAGCCGTCACCGAGCTGCTGCTCGGCGAGTGGCGGCGCGCGCTCTGCTGCGCAAC 63084
QY 388 GCCATGCGCGGCGGAGAGTGTGGGAGCACTGTGGCGCGGCGCTTCCGCTGATCCCGGAC 447

Db 63083 GCCATCGCCGCGCAAGTGTGCGACCTAGTGAGCCCGACCTTCGCTGATCCCGAC 63024

Qy 448 CGCTACCGCGCGAGCGGTCAGCTGTGTCTCCGATGTGAGGCGCGCGCGCG 507

Db 63223 CGCTACCGCGCGAGCGGTCAGCTGTGTCTCCGATGTGAGGCGCGCGCGCG 62864

Qy 508 AAGTGTGCGCTGTGTGCGCTGTGTGCAAGTGTGCAAGGCGCTTCACCAACACTGCG 567

Db 62963 AAGTGTGCGCTGTGTGCGCTGTGTGCAAGTGTGCAAGGCGCTTCACCAACACTGCG 62904

Qy 568 GAGCTCAAGAGCTTCGAGGACCGAGGCGGCTGTGTGCGGCAAGGCGCGGAGCGCGCGCC 627

Db 62903 GAGCTCAAGAGCTTCGAGGACCGAGGCGGCTGTGTGCGGCAAGGCGCGGAGCGCGCC 62844

Qy 628 CGCGCCGCGAGCGCGCAAGCCAAACAGGCGGCTGTGTGAGAACGCTTCTAGAGCGCGCC 687

Db 62843 CGCGCCGCGAGCGCGCAAGCCAAACAGGCGGCTGTGTGAGAACGCTTCTAGAGCGCGCC 62784

Qy 688 GCGGCTTCTTCCCAACCGCGCGCGCGCGCGCGCTGTGTGACCGCGCGCCCACTTCTGTCTTC 747

Db 62783 GCGGCTTCTTCCCAACCGCGCGCGCGCGCGCGCTGTGTGACCGCGCGCCCACTTCTGTCTTC 62724

Qy 748 TGCGCTGTGT 759

Db 62723 TGCGCTGTGT 62712

RESULT 11
AC003098/c 94752 bp DNA linear PRI 29-JAN-1998
LOCUS Homo sapiens chromosome 17, clone HRPc905N1, complete sequence.
DEFINITION AC003098
AC003098 GI:2822155
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS

1 (bases 1 to 94752)
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
Birren, B., Fasman, K., McKernan, K., Nusbaum, C., Richardson, P. and Lander, E.

2 (bases 1 to 94752)
Homo sapiens chromosome 17, clone HRPc905N1
Birren, B., Fasman, K., McKernan, K., Munro, C., Nusbaum, C., Richardson, P., Lander, E., Baldwin, J., Barna, N., Cantu, C., Chang, A., Cooke, P., Daly, M. J., Devon, K., Dewar, K., Dubette, B., Forrest, C., Gage, D., Gensheimer, S., Geraghty, K., Gilmartin, T., Hagos, B., Halphen, I., Harris, K., Howland, J. C., Huang, U., Hui, L., Jacotot, L., Klitz, A., Lane, M., Mackenzie, J., Margulis, N., McDermott, J., Molla, M., Morrow, J., Nachman, A., Naylor, J., O'Connor, T., Oloru, A., Peterson, K., Roberts, D., Rollins, G., Sarnak, A., Shiu, P., Shyam, R., Stillwell, J., Stone, C., Strickland, C., Sydney, K., Tang, L., Zemseva, I. and Zody, M.

3 (bases 1 to 94752)
Submitted (10-NOV-1997) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
Birren, B., Fasman, K., McKernan, K., Nusbaum, C., Richardson, P., Lander, E., Allen, N., Baker, J., Baldwin, J., Barna, N., Beckery, R., Bouwell, C., Byrne, S., Cantu, C., Castle, A., Cooke, P., Daly, M. J., Depayre, B., Devon, K., Dewar, K., Dubette, B., Elmadi, S., Ferreira, P., Forrest, C., Gage, D., Gardyna, S., Gensheimer, S., Geraghty, K., Gilmartin, T., Gray, D., Hagos, B., Harris, K., Horton, L., Howland, J. C., Hui, L., Jacotot, L., Linton, L., Mackenzie, J., Margulis, N., McEwan, P., McGarr, A., Meldrum, J., Molla, M., Morris, W., Morrow, J., Nachman, A., Naylor, J., O'Connor, T., Pavlin, B., Peterson, K., Ranganath, S., Riley, R., Roberts, D., Rollins, G., Rossello, R., Roy, A., Shyam, R., Sotho, S., Stillwell, J., Stone, C., Strickland, C., Sydney, K., Tang, L., Vassiliou, H., Vo, A., Wagner, A., Wheeler, J., Wu, Y., Ye, W. J., Zemseva, I., Zhao, J. and Zody, M.

Direct Submission
Submitted (10-NOV-1997) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 94752)

TITLE
JOURNAL
REFERENCE
AUTHORS

JOURNAL
COMMENT

Submitted (29-JAN-1998) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jan 29, 1998 this sequence version replaced gi:2665504.
The Staden database, finishing information, and all
chromatographic files used in the assembly of this clone are
available from our anonymous ftp site.

All repeats were identified using RepeatMasker: Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>
Location/Qualifiers

FEATURES
source

1. 94752
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="17"
/map="17"
/clone="HRPc905N1"
/clone_lib="Peter de Jong/ human PAC library"
86. .439
/rpt_family="LINE2"
complement(972. .1262)
/rpt_family="AluX"
complement(1884. .1912)
/rpt_family="CA)n"
1912. .1959
/rpt_family="CA)n"
2633. .2681
/rpt_family="MIR"
2701. .3002
/rpt_family="AluX"
complement(3874. .4008)
/rpt_family="LINE2"
4513. .4543
/rpt_family="AT_rich"
complement(5202. .5237)
/rpt_family="GA)n"
complement(5310. .5406)
/rpt_family="MIR"
complement(5829. .5868)
/rpt_family="GC_rich"
6485. .6592
/rpt_family="MIR"
complement(6973. .7029)
/rpt_family="MIR"
complement(10402. .10426)
/rpt_family="GC_rich"
complement(11421. .11496)
/rpt_family="LINE"
complement(11505. .11804)
/rpt_family="AluX"
11811. .11879
/rpt_family="AT_rich"
11880. .12093
/rpt_family="AluX"
complement(12104. .12476)
/rpt_family="LINE2"
12478. .12773
/rpt_family="AluX"
complement(12775. .12917)
/rpt_family="LINE2"
complement(13088. .13510)
/rpt_family="MLT1C"
complement(13544. .13847)
/rpt_family="AluX"
complement(14110. .14776)
/rpt_family="LINE1"
complement(14777. .15077)
/rpt_family="AluX"
complement(15079. .15282)
/rpt_family="LINE1"
15954. .16024
/rpt_family="AT_rich"
18031. .18178

```

/rpt_family="MIR"
repeat_region 20398..20769
/rpt_family="LINE"
repeat_region 20893..21190
/rpt_family="AluSg"
repeat_region complement(21275..21473)
/rpt_family="AluSg"
repeat_region complement(21485..21556)
/rpt_family="(CA)n"
repeat_region 21567..21906
/rpt_family="MER2"
repeat_region 21907..21952
/rpt_family="MLT2CB"
repeat_region 21974..22115
/rpt_family="MIR"
repeat_region 22238..22540
/rpt_family="AluSx"
repeat_region 22577..22633
/rpt_family="LINE2"
repeat_region complement(23190..23270)
/rpt_family="LINE2"
repeat_region 24194..24324
/rpt_family="FLAM_A"
repeat_region 24329..24643
/rpt_family="AluSp"
repeat_region 25542..25698
/rpt_family="LINE2"
repeat_region 25805..26001
/rpt_family="MIR"
repeat_region 26060..26174
/rpt_family="LINE2"
repeat_region 26343..26649
/rpt_family="AluSx"
repeat_region complement(27491..27783)
/rpt_family="AluSg"
repeat_region 27862..27906
/rpt_family="LINE2"
repeat_region 28068..28114
/rpt_family="MIR"
repeat_region complement(28329..28360)
/rpt_family="AT_rich"
repeat_region 28371..28652
/rpt_family="AluSg"
repeat_region complement(29498..29604)
/rpt_family="GC_rich"
repeat_region 30074..30181
/rpt_family="MIR"
repeat_region complement(30181..30249)
/rpt_family="LINE2"
repeat_region 30444..30517
/rpt_family="MIR"
repeat_region complement(31852..32135)
/rpt_family="AluSg"
repeat_region complement(32675..33003)
/rpt_family="MLT1A1"
repeat_region complement(33042..33082)
/rpt_family="LINE2"
repeat_region 33363..33459
/rpt_family="MER5B"
repeat_region complement(35475..35524)
/rpt_family="MLT2D"
repeat_region 35495..35591
/rpt_family="(GA)n"
repeat_region complement(35561..35673)
/rpt_family="MLT2D"
repeat_region complement(35678..35935)
/rpt_family="MER6B"
repeat_region complement(35936..36233)
/rpt_family="AluSg"
repeat_region complement(36236..36449)
/rpt_family="MER6B"
repeat_region complement(36450..36867)
/rpt_family="MLT2D"

```

```

repeat_region complement(36876..37460)
/rpt_family="LINE2"
repeat_region complement(37477..37777)
/rpt_family="AluSx"
repeat_region complement(37804..37948)
/rpt_family="L1"
repeat_region 37948..38130
/rpt_family="LINE1"
repeat_region 38175..38253
/rpt_family="LINE2"

```

Query Match 68.1%; Score 516.8; DB 9; Length 94752;
 Best Local Similarity 96.0%; Pred. No. 4.8e-70;
 Matches 530; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

```

QY 208 CCGGCGAGAAACGAGGAGCGGCTCCCAACACCCCTTTGAGACCAAGAAGTGTCCGAG 267
DB 6340 CAGGGGTGAGCAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 6281
QY 268 TACAGCTGCGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 327
DB 6280 TACAGCTGCGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6221
QY 328 AAGCGGTGACGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 387
DB 6220 AAGCGGTGACGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6161
QY 388 GCCATGCGCGCGGCGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 447
DB 6160 GCCATGCGCGCGGCGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6101
QY 448 CGTACCGCGCGGCGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 507
DB 6100 CGTACCGCGCGGCGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6041
QY 508 AAGTGTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 567
DB 6040 AAGTGTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5991
QY 568 GAGCTCAAGAGCTTGGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 627
DB 5980 GAGCTCAAGAGCTTGGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 5921
QY 628 CGCGCGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 687
DB 5920 CGCGCGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 5861
QY 688 GGGCGGCTGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 747
DB 5860 GGGCGGCTGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 5801
QY 748 TGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 759
DB 5800 TGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5789

```

RESULT 12
 AF326740 638 bp mRNA linear ROD 28-FEB-2001
 LOCUS AF326740
 DEFINITION Mus musculus sclerostin mRNA, complete cds.
 ACCESSION AF326740
 VERSION AF326740.1 GI:13161022
 KEYWORDS
 SOURCE
 ORGANISM
 Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
 AUTHORS
 1 (bases 1 to 638)
 Brunkow, M.E., Gardner, J.C., Van Ness, J., Paepker, B.W.,
 Kovacevich, B.R., Prohl, S., Skonier, J.E., Zhao, L., Sabo, P.J.,
 Fu, Y.H., Altsch, R.S., Gillet, L., Colbert, T., Taccanti, P., Galas, D.,
 Hamersma, H., Beighton, P. and Mulligan, J.T.
 Bone dysplasia sclerostosis results from loss of the SOST gene

QY 637 AGCGCCAAAGCCACGAGCGCGAGTGGAGAAAGCCTTACTAG 678
 DB 595 GGAGGCCAAAGCCACGAGCGAGTGGAGAAAGCCTTACTAG 636
 RESULT 14
 AF326738 532 bp mRNA linear MAM 28-FEB-2001
 LOCUS Bos taurus sclerostin mRNA, partial cds.
 DEFINITION AF326738
 ACCESSION AF326738.1 GI:13161016
 VERSION
 KEYWORDS
 ORGANISM Bos taurus.
 Bos taurus.
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 Bovidae; Bovinae; Bos.
 REFERENCE 1 (bases 1 to 532)
 Brunkow, M.E., Gardner, J.C., Van Ness, J., Paepfer, B.W.,
 Kovacevich, B.R., Prohl, S., Skonier, J.E., Zhao, L., Sabo, P.J.,
 Fu, Y.H., Alisch, R.S., Gillett, L., Colbert, T., Tacconi, P., Galas, D.,
 Hamersma, H., Beighton, P. and Mulligan, J.T.
 Bone dysplasia sclerostosis results from loss of the SOST gene
 product, a novel cysteine knot-containing protein
 Am. J. Hum. Genet. 68 (3), 577-589 (2001)
 JOURNAL MEDLINE 21090529
 PUBMED 11179006
 REFERENCE 2 (bases 1 to 532)
 Brunkow, M.E., Gardner, J.C., Van Ness, J., Paepfer, B.W.,
 Kovacevich, B.R., Prohl, S., Skonier, J.E., Zhao, L., Sabo, P.J.,
 Fu, Y.H., Alisch, R.S., Gillett, L., Colbert, T., Tacconi, P., Galas, D.,
 Hamersma, H., Beighton, P. and Mulligan, J.T.
 Direct Submission
 Submitted (07-DEC-2000) Genomics, Celltech Chiroscience Inc., 1631
 220th St. SE, Bothell, WA 98021, USA
 JOURNAL TITLE Location/Qualifiers
 FEATURES
 source 1..532
 /organism="Bos taurus"
 /db_xref="taxon:9913"
 CDS <1..>532
 /product="sclerostin"
 /codon_start=3
 /protein_id="AAK13453.1"
 /db_xref="GI:13161017"
 /translation="NDATRTITPLGTYPPPLPINKTKMRAENGRRPHHPFTKDA
 SEYSRELFHRTVTDGRCRSAPVTEIVCSQCGPARLTPVAIGRGKMKRSGDFR
 CIPDRYARQRVOLCPGGAAPRARKVRLVASCCKRLTRFHNSLKKDFGPAARPQT
 GRKLRPRAGTKASRA"
 BASE COUNT 95 a 199 c 176 g 62 t
 ORIGIN
 Query Match 61.6%; Score 467.8; DB 4; Length 532;
 Best Local Similarity 93.5%; Pred. No. 4.3e-62;
 Matches 500; Conservative 0; Mismatches 32; Indels 3; Gaps 1;
 QY 125 AGATGATGCGGAGGAGTATCCCGAGCTCGGAGGAGTACCCCGAGCTCCACCGGAGC 184
 DB 1 AGATGATGCGGAGGAGTATCCCGAGCTCGGAGGAGTACCCCGAGCTCCACCGGAGC 60
 QY 185 TGGAGAACAAACAGACCATGAACCGGCGGAGAGAACGAGGCGGCTCCACCGACCCCT 244
 DB 61 T--GAACAAACAGACCATGAACCGGCGGAGAGAACGAGGAGACCTCCACCGACCCCT 117
 QY 245 TTGAGACCAAGAGCGTGTCCAGTACAGCTCCCGGAGCTGCATTTACCCCGTCACTGA 304
 DB 118 TTGAGACCAAGAGCGTGTCCAGTACAGCTCCCGGAGCTGCATTTACCCCGTCACTGA 177
 QY 305 CCGATGGCGCGTGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 364
 DB 178 CCGATGGCGCGTGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 237
 QY 365 GCCGGCGCGCTGCTGCCCAAGCGCATCGGCGGCGGCAAGTGTGGCGACTTGTGAGG 424

DB 238 GCCCGGCGCGCTGCTGCCCAAGCCATCGGCGGCGGCAAGTGTGGCGGCCCAAGCGGCGC 297
 QY 425 CCGACTTCGCTGATTCCTCCGAGACCGCTACCGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 484
 DB 298 CCGACTTCGCTGATTCCTCCGAGACCGCTACCGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 357
 QY 485 GTGTGAGGCGCGCGGCGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 544
 DB 358 GCGGCGGCGCGCGGCGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 417
 QY 545 TCACCGCTTCCAAACAGTGTGAGTCAAGGACTTTGGGAGCCGAGCGGCGGCGGCGGCGG 604
 DB 418 TCACCTGCTTCCAAACAGTGTGAGTCAAGGACTTTGGGAGCGGAGCGGAGCGGAGCGGAGCGG 477
 QY 605 AGAAGGCGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 659
 DB 478 AAGGCGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 532
 RESULT 15
 AC073954 177744 bp DNA linear HTG 17-NOV-2000
 LOCUS Homo sapiens chromosome 17 clone RP11-147C10, WORKING DRAFT
 DEFINITION AC073954
 ACCESSION AC073954.3 GI:10988993
 VERSION
 KEYWORDS HTG; HTGS PHASE1; HTGS_DRAFT.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 177744)
 Waterston, R.H.
 The sequence of Homo sapiens clone
 Unpublished
 2 (bases 1 to 177744)
 Waterston, R.H.
 Direct Submission
 Submitted (07-JUL-2000) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA
 On Oct 25, 2000 this sequence version replaced gi:9186849.
 COMMENT
 ----- Genome Center -----
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: http://genome.wustl.edu/gsc/index.shtml
 Project Information -----
 Center project name: H_NH014/C10
 ----- Summary Statistics -----
 Sequencing vector: M13; 100%
 Sequencing method: Dye-terminator Big Dye; 0% of reads
 Chemistry: Dye-terminator Big Dye; 0% of reads
 Assembly program: Phrap; version 0.99019
 Consensus quality: 155679 bases at least Q40
 Consensus quality: 163329 bases at least Q30
 Consensus quality: 166780 bases at least Q20
 Insert size: 185000; agarose-fp
 Insert size: 183559; sum-of-contigs
 Quality coverage: 3.29 in Q20 bases; sum-of-contigs
 Quality coverage: 3.43 in Q20 bases; sum-of-contigs
 ----- NOTE: This is a 'working draft' sequence. It currently
 * consists of 34 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 * 1 1247: contig of 1247 bp in length
 * 1248 1347: gap of unknown length

```

* 1348 2762: contig of 1415 bp in length
* 2763 2862: gap of unknown length
* 2863 4773: contig of 1911 bp in length
* 4774 4873: gap of unknown length
* 4874 6326: contig of 1453 bp in length
* 6327 6426: gap of unknown length
* 6427 8873: contig of 2447 bp in length
* 8874 8974: gap of unknown length
* 8974 10783: contig of 1810 bp in length
* 10784 10883: gap of unknown length
* 10884 12763: contig of 1880 bp in length
* 12764 12863: gap of unknown length
* 12864 15381: contig of 2518 bp in length
* 15382 15481: gap of unknown length
* 15482 17365: contig of 1884 bp in length
* 17366 17465: gap of unknown length
* 17466 19993: contig of 2527 bp in length
* 19993 20092: gap of unknown length
* 20093 23512: contig of 3420 bp in length
* 23513 23612: gap of unknown length
* 23613 25699: contig of 2087 bp in length
* 25700 25799: gap of unknown length
* 25800 28333: contig of 2534 bp in length
* 28334 28433: gap of unknown length
* 28434 32027: contig of 3594 bp in length
* 32028 32127: gap of unknown length
* 32128 36115: contig of 3988 bp in length
* 36116 36215: gap of unknown length
* 36216 39882: contig of 3667 bp in length
* 39883 39983: gap of unknown length
* 39984 44009: contig of 4027 bp in length
* 44010 44109: gap of unknown length
* 44110 47829: contig of 3720 bp in length
* 47830 47929: gap of unknown length
* 47930 53768: contig of 5839 bp in length
* 53769 53868: gap of unknown length
* 53869 60215: contig of 6347 bp in length
* 60216 60315: gap of unknown length
* 60316 66477: contig of 6162 bp in length
* 66478 66577: gap of unknown length
* 66578 71476: contig of 4899 bp in length
* 71477 71576: gap of unknown length
* 71577 78883: contig of 7307 bp in length
* 78884 78983: gap of unknown length
* 78984 85952: contig of 6969 bp in length
* 85953 86052: gap of unknown length
* 86053 93510: contig of 7458 bp in length
* 93511 93610: gap of unknown length
* 93611 100718: contig of 7108 bp in length
* 100719 100818: gap of unknown length
* 100819 109795: contig of 8977 bp in length
* 109796 109895: gap of unknown length
* 109896 117790: contig of 7895 bp in length
* 117791 117890: gap of unknown length
* 117891 126044: contig of 8154 bp in length
* 126045 126144: gap of unknown length
* 126145 134442: contig of 8298 bp in length
* 134443 134542: gap of unknown length
* 134543 145663: contig of 11121 bp in length
* 145664 145763: gap of unknown length
* 145764 157189: contig of 11426 bp in length
* 157190 157289: gap of unknown length
* 157290 166822: contig of 9533 bp in length
* 166823 177744: contig of 10822 bp in length.

```

FEATURES

```

source
1.177744
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="17"
/clone="RP11-147C10"
1..1247
/note="assembly_name:Contig26"
misc_feature

```

```

misc_feature 1348..2762
/note="assembly_name:Contig31"
misc_feature 2863..4773
/note="assembly_name:Contig34"
misc_feature 4874..6326
/note="assembly_name:Contig35"
misc_feature 6427..8873
/note="assembly_name:Contig36"
misc_feature 8974..10783
/note="assembly_name:Contig37"
misc_feature 10884..12763
/note="assembly_name:Contig38"
misc_feature 12864..15381
/note="assembly_name:Contig39"
misc_feature 15482..17365
/note="assembly_name:Contig40"
misc_feature 17466..19992
/note="assembly_name:Contig41"
misc_feature 20093..23512
/note="assembly_name:Contig42"
misc_feature 23613..25699
/note="assembly_name:Contig43"
misc_feature 25800..28333
/note="assembly_name:Contig44"
misc_feature 28434..32027
/note="assembly_name:Contig45"
misc_feature 32128..36115
/note="assembly_name:Contig46"
misc_feature 36216..39882
/note="assembly_name:Contig47"
misc_feature 39983..44009
/note="assembly_name:Contig48"
misc_feature 44110..47829
/note="assembly_name:Contig49"
misc_feature 47930..53768
/note="assembly_name:Contig50"
misc_feature 53869..60215
/note="assembly_name:Contig51"
misc_feature 60316..66477
/note="assembly_name:Contig52"
misc_feature 66578..71476
/note="assembly_name:Contig53"
misc_feature 71577..78883
/note="assembly_name:Contig54"
misc_feature 78984..85952
/note="assembly_name:Contig55"
misc_feature 86053..93510
vector_side:left"
misc_feature 93611..100718
/note="assembly_name:Contig56"
misc_feature 100819..109795
/note="assembly_name:Contig57"
misc_feature 109896..117790
/note="assembly_name:Contig58"
misc_feature 117891..126044
/note="assembly_name:Contig59"
misc_feature 126145..134442
/note="assembly_name:Contig60"
misc_feature 134543..145663
/note="assembly_name:Contig61"
misc_feature 145764..157189
/note="assembly_name:Contig62"
misc_feature 157290..166822
/note="assembly_name:Contig63"
misc_feature 166923..177744
/note="assembly_name:Contig64"
misc_feature 177745..188888
/note="assembly_name:Contig65"

```

ORIGIN

```

BASE COUNT 44147 a 43012 c 43431 g 43839 t 3315 others
Query Match 54.9%; Score 417; DB 2; Length 177744;
Best Local Similarity 91.0%; Pred. No. 6.ee-55;

```


GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 28, 2003, 17:25:48 ; Search time 2063.01 Seconds

(without alignments)
8972.025 Million cell updates/sec

Title: US-09-867-274-3

Perfect score: 636
Sequence: 1 atgcagccctcactagcccc.....agctgagagcctcactag 636

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

GeneBml:*

1: gb_ba:*

2: gb_hcg:*

3: gb_in:*

4: gb_om:*

5: gb_ov:*

6: gb_pac:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_scs:*

12: gb_sy:*

13: gb_un:*

14: gb_vl:*

15: em_ba:*

16: em_fun:*

17: em_hum:*

18: em_in:*

19: em_mu:*

20: em_om:*

21: em_or:*

22: em_ov:*

23: em_pat:*

24: em_ph:*

25: em_pl:*

26: em_ro:*

27: em_scs:*

28: em_un:*

29: em_vl:*

30: em_hcg_hum:*

31: em_hcg_inv:*

32: em_hcg_other:*

33: em_hcg_mus:*

34: em_hcg_pln:*

35: em_hcg_rod:*

36: em_hcg_mam:*

37: em_hcg_vtc:*

38: em_sy:*

39: em_hcgo_hum:*

40: em_hcgo_mus:*

41: em_hcgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	636	100.0	636	6 AX323455	AX323455 Sequence
2	634.4	99.7	638	10 AF326740	AF326740 Mus muscu
3	557.6	87.7	674	10 AF326741	AF326741 Rattus no
4	504.8	79.4	642	9 AF326742	AF326742 Cercopit
5	501.6	78.9	759	6 AX323453	AX323453 Sequence
6	501.6	78.9	2271	6 AX342535	AX342535 Sequence
7	501.6	78.9	2296	9 AF331844	AF331844 Homo sapi
8	501.6	78.9	2323	9 AF326739	AF326739 Homo sapi
9	501.6	78.9	2329	6 AX056687	AX056687 Sequence
10	426.6	67.1	81806	10 AF326737	AF326737 Mus muscu
11	426.6	67.1	110000	2 AC068782-2	Continuation (3 of
12	426.6	67.1	205277	2 AC012296	AC012296 Mus muscu
13	426.6	67.1	208135	10 AC068807	AC068807 Mus muscu
14	425	66.8	198508	10 AL591145	AL591145 Mouse DNA
15	415.8	65.4	532	4 AF326738	AF326738 Bos tauru
16	409	64.3	101804	2 AC098160	AC098160 Rattus no
17	409	64.3	104898	2 AC121721	AC121721 Rattus no
18	367	57.7	7099	6 AX342537	AX342537 Sequence
19	367	57.7	21501	9 AF326736	AF326736 Homo sapi
20	367	57.7	80117	9 AC055813	AC055813 Homo sapi
21	367	57.7	94752	9 AC003098	AC003098 Homo sapi
22	282.6	44.4	177744	2 AC073954	AC073954 Homo sapi
23	214.4	33.7	51575	2 AC023810	AC023810 Mus muscu
24	83.2	13.1	968	10 BC021458	BC021458 Mus muscu
25	81.2	12.8	220967	2 AC124776	AC124776 Mus muscu
26	71.6	11.3	182253	2 AC117353	AC117353 Rattus no
27	67.2	10.6	900	6 AF017832	AF017832 Sequence
28	67.2	10.6	1091	6 AX454484	AX454484 Sequence
29	67.2	10.6	1091	6 AX490962	AX490962 Sequence
30	67.2	10.6	1723	6 BC008484	BC008484 Homo sapi
31	65.6	10.3	618	6 AX429979	AX429979 Sequence
32	65.6	10.3	621	6 AX429978	AX429978 Sequence
33	65.6	10.3	1692	6 AX429953	AX429953 Sequence
34	65.2	10.3	2409	6 AK093408	AK093408 Homo sapi
35	65.2	10.3	125372	9 AC079155	AC079155 Homo sapi
36	63.6	10.0	1500	9 HSM800108	AL050024 Homo sapi
37	59.2	9.3	32748	1 AB070951	AB070951 Streptomy
38	59.2	9.3	37638	2 AC111799	AC111799 Rattus no
39	59	9.3	78862	2 AC095361	AC095361 Rattus no
40	57.4	9.0	172490	2 AC117097	AC117097 Rattus no
41	57	9.0	412	6 AX071502	AX071502 Sequence
42	55.2	8.7	125020	9 AF429315	AF429315 Homo sapi
43	54.6	8.6	92055	2 AC105692	AC105692 Rattus no
44	54.4	8.6	132884	2 AC128403	AC128403 Rattus no
45	54.4	8.6	214885	2 AC094318	AC094318 Rattus no

ALIGNMENTS

RESULT 1
AX323455
LOCUS AX323455
DEFINITION Sequence 3 from Patent WO01923308.
ACCESSION AX323455
VERSION AX323455.1 GI:18094217
KEYWORDS
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS Paszty, C. J. and Gao, Y.
TITLE Cytidine-knot polypeptides: cloaked-2 molecules and uses thereof
JOURNAL Patent: WO 01923308-A 3 06-DEC-2001;

Amgen, Inc. (US)
 FEATURES Location/Qualifiers
 source 1..636
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 BASE COUNT 114 a 224 c 207 g 91 t
 ORIGIN

Query Match 100.0%; Score 636; DB 6; Length 636;
 Best Local Similarity 100.0%; Pred. No. 4.8e-99;
 Matches 636; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGAGGCTCTACATGAGCCGCTGCTCATCTGCTCTATGTCAGCGCTTCTGCT 60
 1 ATGAGGCTCTACATGAGCCGCTGCTCATCTGCTCTATGTCAGCGCTTCTGCT 60
 61 GTGAGGGGCGAGGGGTGGCAAGCTTTCAGAGATGATCCAGAGGTCTATCCAGGGCTT 120
 61 GTGAGGGGCGAGGGGTGGCAAGCTTTCAGAGATGATCCAGAGGTCTATCCAGGGCTT 120
 121 GGAAGATACCCCGAGGCTCTCTGAGAACCAACGACATGAAACCGGGCGAGATGA 180
 121 GGAAGATACCCCGAGGCTCTCTGAGAACCAACGACATGAAACCGGGCGAGATGA 180
 181 GGCAGACCTCCCGACATCCCTATGAGCCCAAGATGTGTCCAGATGACAGTCCGCGAG 240
 181 GGCAGACCTCCCGACATCCCTATGAGCCCAAGATGTGTCCAGATGACAGTCCGCGAG 240
 241 CTGCACTACACCCGCTTCTGAGAACGAGCCCATGCGCAGCGCGCAAGCGGTCACCGAG 300
 241 CTGCACTACACCCGCTTCTGAGAACGAGCCCATGCGCAGCGCGCAAGCGGTCACCGAG 300
 301 TTGGTGTGCTCCCGGCGAGTGGCGCCCGCGCGCTGCTGCGCAACCGCATCCGGCGGTG 360
 301 TTGGTGTGCTCCCGGCGAGTGGCGCCCGCGCGCTGCTGCGCAACCGCATCCGGCGGTG 360
 361 AAGTGTGGCGCGCGAGCGAGCCGATTTCCGCTGATCCCGATCCCGATCCCGCGAG 420
 361 AAGTGTGGCGCGCGAGCGAGCCGATTTCCGCTGATCCCGATCCCGATCCCGCGAG 420
 421 CGGGTGCAGCTGCTGTGCGCGCGCGCGCGCGCGCTGCGCAAGGTGCGTCTGCTG 480
 421 CGGGTGCAGCTGCTGTGCGCGCGCGCGCGCGCGCTGCGCAAGGTGCGTCTGCTG 480
 481 GCTCTGTGCAAGTGCAGAGCGCTTCCAGAACCAAGTGCAGAGCTCAAGGACTTC 540
 481 GCTCTGTGCAAGTGCAGAGCGCTTCCAGAACCAAGTGCAGAGCTCAAGGACTTC 540
 541 GGGCGGAGACCGCGCGCGCGCGAGAGGGTGCAGAGCGCGCGCGCGCGCGCGAGCC 600
 541 GGGCGGAGACCGCGCGCGCGCGAGAGGGTGCAGAGCGCGCGCGCGCGCGCGAGCC 600
 601 AAAGCCAAACGAGCGAGCTGAGAACGCTTACTAG 636
 601 AAAGCCAAACGAGCGAGCTGAGAACGCTTACTAG 636

RESULT 2
 AF326740 638 bp mRNA linear ROD 28-FEB-2001
 LOCUS AF326740
 DEFINITION Mus musculus sclerostin mRNA, complete cds.
 ACCESSION AF326740
 VERSION AF326740.1 GI:13161022
 KEYWORDS
 SOURCE Mus musculus.
 ORGANISM Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 638)
 Brunkow,M.E., Gardner,J.C., Van Ness,J., Paepel,B.W.,
 Kovacevich,B.R., Prohl,S., Skonier,J.E., Zhao,L., Sabo,P.J.,
 Fu,Y.H., Alisch,R.S., Gillett,L., Colbert,T., Tacconi,P., Galas,D.,
 Hamerema,H., Beighton,P. and Mulligan,J.T.

TITLE Bone dysplasia sclerosteosis results from loss of the SOST gene
 JOURNAL Am. J. Hum. Genet. 68 (3), 577-589 (2001)
 MEDLINE 21090529
 PUBMED 11179006
 REFERENCE 2 (bases 1 to 638)
 AUTHORS Brunkow,M.E., Gardner,J.C., Van Ness,J., Paepel,B.W.,
 Kovacevich,B.R., Prohl,S., Skonier,J.E., Zhao,L., Sabo,P.J.,
 Fu,Y.H., Alisch,R.S., Gillett,L., Colbert,T., Tacconi,P., Galas,D.,
 Hamerema,H., Beighton,P. and Mulligan,J.T.
 Direct Submission
 Submitted (07-DEC-2000) Genomics, Celltech Chiroscience Inc., 1631
 220th St. SE, Bothell, WA 98021, USA

FEATURES Location/Qualifiers
 source 1..638
 /organism="Mus musculus"
 /strain="129Sv/J"
 /db_xref="taxon:10090"
 /chromosome="11"
 1..636
 /codon_start=1
 /product="sclerostin"
 /protein_id="P4K13455.1"
 /db_xref="GI:13161023"
 /translation="MQPSLAPCLICLVHNAFCVAGGQGVQAFRNDATVETPGLSEYP
 EPPENNQTMRAENGRPPHHPDAKGVSEYSCRELHYTRFLTDGPRSAKVTLLV
 CGCGCGPRLPLPAIGRWKWRPNRPFRCIPRYAROROLLCPGAAPRSKRVLV
 ASCCKRLTFPHNSELKDPEPTARQKRKRPPARAKNAQALENNY"

BASE COUNT 114 a 224 c 209 g 91 t
 ORIGIN

Query Match 99.7%; Score 634.4; DB 10; Length 638;
 Best Local Similarity 99.8%; Pred. No. 9e-99; 1; Indels 0; Gaps 0;
 Matches 635; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 ATGAGCCCTCACTAGCCCGCTGCTCATCTGCTCTATGTCAGCGCTTCTGCTGCT 60
 1 ATGAGCCCTCACTAGCCCGCTGCTCATCTGCTCTATGTCAGCGCTTCTGCTGCT 60
 61 GTGAGGGGCGAGGGGTGGCAAGCTTTCAGAGATGATCCAGAGGTCTATCCAGGGCTT 120
 61 GTGAGGGGCGAGGGGTGGCAAGCTTTCAGAGATGATCCAGAGGTCTATCCAGGGCTT 120
 121 GGAAGATACCCCGAGGCTCTCTGAGAACCAACGACATGAAACCGGGCGAGATGA 180
 121 GGAAGATACCCCGAGGCTCTCTGAGAACCAACGACATGAAACCGGGCGAGATGA 180
 181 GGCAGACCTCCCGACATCCCTATGAGCCCAAGATGTGTCCAGATGACAGTCCGCGAG 240
 181 GGCAGACCTCCCGACATCCCTATGAGCCCAAGATGTGTCCAGATGACAGTCCGCGAG 240
 241 CTGCACTACACCCGCTTCTGAGAACGAGCCCATGCGCAGCGCGCAAGCGGTCACCGAG 300
 241 CTGCACTACACCCGCTTCTGAGAACGAGCCCATGCGCAGCGCGCAAGCGGTCACCGAG 300
 301 TTGGTGTGCTCCCGGCGAGTGGCGCCCGCGCGCTGCTGCGCAACCGCATCCGGCGGTG 360
 301 TTGGTGTGCTCCCGGCGAGTGGCGCCCGCGCGCTGCTGCGCAACCGCATCCGGCGGTG 360
 361 AAGTGTGGCGCGCGAGCGAGCCGATTTCCGCTGATCCCGATCCCGATCCCGCGAG 420
 361 AAGTGTGGCGCGCGAGCGAGCCGATTTCCGCTGATCCCGATCCCGATCCCGCGAG 420
 421 CGGGTGCAGCTGCTGTGCGCGCGCGCGCGCGCGCTGCGCAAGGTGCGTCTGCTG 480
 421 CGGGTGCAGCTGCTGTGCGCGCGCGCGCGCGCGCTGCGCAAGGTGCGTCTGCTG 480
 481 GCTCTGTGCAAGTGCAGAGCGCTTCCAGAACCAAGTGCAGAGCTCAAGGACTTC 540
 481 GCTCTGTGCAAGTGCAGAGCGCTTCCAGAACCAAGTGCAGAGCTCAAGGACTTC 540
 541 GGGCGGAGACCGCGCGCGCGCGAGAGGGTGCAGAGCGCGCGCGCGCGCGCGAGCC 600
 541 GGGCGGAGACCGCGCGCGCGCGAGAGGGTGCAGAGCGCGCGCGCGCGCGCGAGCC 600

Db 541 GGGCCGAGACCGCGCGCCGAGAGGTCGCAAGCCGCGCCCGCGCGCGCGAGCC 600

Qy 601 AAAGCAACGAGCGAGCTTGAGAGCGCTTACTAG 636

Db 601 AAAGCAACGAGCGAGCTTGAGAGCGCTTACTAG 636

RESULT 3
AF326741

LOCUS
Rattus norvegicus scleroetin mRNA, complete cds.

DEFINITION
AF326741

ACCESSION
AF326741.1 GI:13161025

VERSION
AF326741.1

KEYWORDS
Rattus norvegicus.

SOURCE
Rattus norvegicus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE
1 (bases 1 to 674)
Brunkow, M.E., Gardner, J.C., Van Ness, J., Paepers, B.W.,
Kovacevich, B.R., Proli, S., Skonier, J.E., Zhao, L., Sabo, P.J.,
Fu, Y.H., Alisch, R.S., Gillett, L., Colbert, T., Tacconi, P., Galas, D.,
Hamersma, H., Beighton, P. and Mulligan, J.T.,
Bone dysplasia scleroetosis results from loss of the SOST gene
product, a novel cysteine knot-containing protein
Am. J. Hum. Genet. 68 (3), 577-589 (2001)

JOURNAL
MEDLINE
PUBMED
21090529

REFERENCE
2 (bases 1 to 674)
Brunkow, M.E., Gardner, J.C., Van Ness, J., Paepers, B.W.,
Kovacevich, B.R., Proli, S., Skonier, J.E., Zhao, L., Sabo, P.J.,
Fu, Y.H., Alisch, R.S., Gillett, L., Colbert, T., Tacconi, P., Galas, D.,
Hamersma, H., Beighton, P. and Mulligan, J.T.,
Direct Submission
Submitted (07-DEC-2000) Genomics, Celltech Chiroscience Inc., 1631
220th St. SE, Bothell, WA 98021, USA

TITLE
JOURNAL
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

FEATURES
source
1..674
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
1..32
33..674
/codon_start=1
/product="scleroetin"
/protein_id="AAK13456.1"
/db_xref="GI:13161026"
/translation="MQLSLAPCLACLVHAFVAVESQGMQAFKNDATETIIPGLREYP
EPPELENNKTMRRAENGPRPHNPRTDQVSEYSGREIHTYRFTYDGPGRSAKPYTE
LVSGGCGPARLLPNAIGRKWMRPSPGPRFCIPRYRAQVOLLCPGGAAPRKYR
LVASCCCKRLTFHNSLEKDFGPETARQKRPRPARAGAKNAQAELENAV"

BASE COUNT
128 a 240 c 208 g 98 t

ORIGIN

Query Match 87.7%; Score 557.6; DB 10; Length 674;
Best Local Similarity 93.0%; Pred. No. 1.1e-85;
Matches 557; Conservative 0; Mismatch 39; Indels 6; Gaps 1;

Qy 1 ATGACAGCCCTCACTAGCCCGCTGCTCATCTGCTACTTGTGACGCGCTCTTGTGCT 60

Db 33 ATGACAGCTCTCACTAGCCCGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 92

Qy 61 GTGGAGGGCGGAGGGGTGGCAAGCTTCAGAGATGATGCCACAGAGGTCTATCCAGGGGCTT 120

Db 93 GTGGAGAGCGGAGGGGTGGCAAGCTTCAGAGATGATGCCACAGAGATCTATCCGGGAGCTC 152

Qy 121 GGAGAGTACCCCGAGGCT 174

Db 153 AGAGAGTACCCCGAGGCT 212

Qy 175 AATGAGGAGAGAGCTTCCCAACATCTCTTGAAGCCAAAGATGTGTCCAGTACAGCTGC 234

Db 213 AACGAGGAGAGACCCCCCAACCATCTTATGACACCAAGAGGTGTCGAGTACAGCTGC 272

Qy 235 CGGAGCTGATACACCGCGCTTCTGACAGACGCGCCATATGCCAGCGCGCCAGCGGTC 294

Db 273 CGGAGCTGATACACCGCGCTTCTGACAGACGCGCCATATGCCAGCGCGGTC 332

Qy 295 ACCGAGTTGGTGTCTCCGCGCAGTGGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 354

Db 333 ACCGAGTTGGTGTCTCCGCGCAGTGGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 392

Qy 355 CGGTAAGTGTGGTGGCG 414

Db 393 CGGTAAGTGTGGTGGCG 452

Qy 415 GCGCAGCGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 474

Db 453 GCGCAGCGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 512

Qy 475 CTGGTGGCTGTGTCAGAGTGCAGAGCGCTTCAACCGCTTCAACAGTGCAGTCAAG 534

Db 513 CTGGTGGCTGTGTCAGAGTGCAGAGCGCTTCAACCGCTTCAACAGTGCAGTCAAG 572

Qy 535 GACTTGGCGCGCGAGACCG 594

Db 573 GACTTGGCGCGAGTGCAGAGCG 632

Qy 595 GGAGCCAAAGCCCAACGAGCGAGGCTGAGAGCGCTTACTAG 636

Db 633 GGAGCCAAAGCCCAACGAGCGAGGCTGAGAGCGCTTACTAG 674

RESULT 4
AF326742

LOCUS
Cercopithecus aethiops scleroetin mRNA, complete cds.

DEFINITION
AF326742

ACCESSION
AF326742.1 GI:13161028

VERSION
AF326742.1

KEYWORDS
SOURCE
ORGANISM

FEATURES
source
1..642
/organism="Cercopithecus aethiops"
/db_xref="taxon:9534"
1..642
/codon_start=1
/product="scleroetin"
/protein_id="AAK13457.1"
/db_xref="GI:13161029"
/translation="MQLPLALCLVCLLVHAFVAVESQGMQAFKNDATETIIPGLREYP
EPPELENNKTMRRAENGPRPHNPRTDQVSEYSGREIHTYRFTYDGPGRSAKPYTE
LVSGGCGPARLLPNAIGRKWMRPSPGPRFCIPRYRAQVOLLCPGGAAPRKYR

TITLE
JOURNAL
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
1 (bases 1 to 642)
Brunkow, M.E., Gardner, J.C., Van Ness, J., Paepers, B.W.,
Kovacevich, B.R., Proli, S., Skonier, J.E., Zhao, L., Sabo, P.J.,
Fu, Y.H., Alisch, R.S., Gillett, L., Colbert, T., Tacconi, P., Galas, D.,
Hamersma, H., Beighton, P. and Mulligan, J.T.,
Bone dysplasia scleroetosis results from loss of the SOST gene
product, a novel cysteine knot-containing protein
Am. J. Hum. Genet. 68 (3), 577-589 (2001)

JOURNAL
MEDLINE
PUBMED
21090529

REFERENCE
2 (bases 1 to 642)
Brunkow, M.E., Gardner, J.C., Van Ness, J., Paepers, B.W.,
Kovacevich, B.R., Proli, S., Skonier, J.E., Zhao, L., Sabo, P.J.,
Fu, Y.H., Alisch, R.S., Gillett, L., Colbert, T., Tacconi, P., Galas, D.,
Hamersma, H., Beighton, P. and Mulligan, J.T.,
Direct Submission
Submitted (07-DEC-2000) Genomics, Celltech Chiroscience Inc., 1631
220th St. SE, Bothell, WA 98021, USA

BASE COUNT 113 a 228 c 210 g 91 t
 ORIGIN

Query Match 79.4%; Score 504.8; DB 9; Length 642;
 Best Local Similarity 87.9%; Pred. No. 11e-76;
 Matches 564; Conservative 0; Mismatches 72; Indels 6; Gaps 1;

1 ATGACGCTTCACTAGCCCGCTCATCTGCTTGTGCAAGCTTCTGTGCT 60
 1 ATGCACTCCCACTGCGCTGTGTCTGTCTGCTGTAACGACGCTTCTGTGTA 60
 61 GTGAGGCGCAGGGGTGGCAAGCTTCAAGATGATGCAAGAGTATCTCCAGGCTT 120
 61 GTGAGGCGCAGGGGTGGCAAGCTTCAAGATGATGCAAGAGTATCTCCAGGCTT 120
 121 GGAGATTAACCCGAGCTCTCTCC-----TGAGAACACACAGCATGAAACCGGCGAG 174
 121 GGAGATTAACCCGAGCTCTCTCCACGAGCTGAGAACACACAGCATGAAACCGGCGAG 180
 175 AATGAGGAGACGACCTCCCAACCATCTTATGACGCAAAAGATGTGTCCGATGACAGTGC 234
 181 AATGAGGAGGCGGCTCCCAACCATCTTATGAGACCAAGACGATGTCCGATGACAGTGC 240
 235 CGGAGCTGCACTACACCGCTTCTGACAGAGGCTTATGCGGACGCGCAAGCGGCTC 294
 241 CGAGAGCTGCACTACACCGCTTCTGACAGAGGCTTATGCGGACGCGCAAGCGGCTC 300
 295 ACCGAGTGTGTGTCTCCGCGCAGTGCAGGCGCGCGCGCTGCTGCAACGCGCATCGAG 354
 301 ACCGAGTGTGTGTCTCCGCGCAGTGCAGGCGCGCGCGCTGCTGCAACGCGCATCGAG 360
 355 CGCGTGAAGT 414
 361 CGCGGCAAGT 420
 415 GCGCAGGCGGT 474
 421 GCGCAGGCGGT 480
 475 CTGTGTGCTCTGT 534
 481 CTGTGTGCTCTGT 540
 535 GACTTCCGCGCGCGAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 594
 541 GACTTCCGCGCGCGAGCG 600
 595 GGAGCCAAAGCCAAACGAGCGGAGCTGAGAACGCTTACTAG 636
 601 GGAGCCAAAGCCAAATCAGGCGGAGCTGAGAACGCTTACTAG 642

RESULT 5
 AX323453 759 bp DNA linear PART 07-JAN-2002
 LOCUS
 DEFINITION Sequence 1 from Patent W00192308.
 ACCESSION AX323453
 VERSION AX323453.1 GI:18094216
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE
 1 Paszty,C.J. and Gao,Y.
 Cystine-knot polypeptides: cloaked-2 molecules and uses thereof
 JOURNAL Patent: WO 0192308-A 1 06-DEC-2001;
 Amgen, Inc. (US)
 FEATURES
 source Location/Qualifiers
 1..759
 /db_xref="taxon:9606"

BASE COUNT 125 a 282 c 244 g 108 t
 ORIGIN

Query Match 78.9%; Score 501.6; DB 6; Length 759;
 Best Local Similarity 87.5%; Pred. No. 3.8e-76;
 Matches 562; Conservative 0; Mismatches 74; Indels 6; Gaps 1;

1 ATGAGGCTTCACTAGCCCGCTCATCTGCTTGTGCAAGCTTCTGTGCT 60
 37 ATGAGCTCCCACTGCGCTGTGTCTGTCTGTCTGTGTAACGACCTTCTGTGTA 96
 61 GTGAGGCGCAGGGGTGGCAAGCTTCAAGATGATGCAAGAGTATCTCCAGGCTT 120
 97 GTGAGGCGCAGGGGTGGCAAGCTTCAAGATGATGCAAGAGTATCTCCAGGCTT 156
 121 GGAGATTAACCCGAGCTCTCTCC-----TGAGAACACACAGCATGAAACCGGCGAG 174
 157 GGAGATTAACCCGAGCTCTCTCCACGAGCTGAGAACACACAGCATGAAACCGGCGAG 216
 175 AATGAGGAGACGACCTCCCAACCATCTTATGACGCAAAAGATGTGTCCGATGACAGTGC 234
 217 AATGAGGAGGCGGCTCCCAACCATCTTATGAGACCAAGACGATGTCCGATGACAGTGC 276
 235 CGGAGCTGCACTACACCGCTTCTGACAGAGGCTTATGCGGACGCGCAAGCGGCTC 294
 277 CGGAGCTGCACTACACCGCTTCTGACAGAGGCTTATGCGGACGCGCAAGCGGCTC 336
 295 ACCGAGTGTGTGTCTCCGCGCAGTGCAGGCGCGCGCGCTGCTGCAACGCGCATCGAG 354
 337 ACCGAGTGTGTGTCTCCGCGCAGTGCAGGCGCGCGCGCTGCTGCAACGCGCATCGAG 396
 355 CGCGTGAAGT 414
 361 CGCGGCAAGT 420
 415 GCGCAGGCGGT 474
 457 GCGCAGGCGGT 516
 475 CTGTGTGCTCTGT 534
 517 CTGTGTGCTCTGT 576
 535 GACTTCCGCGCGCGAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 594
 577 GACTTCCGCGCGCGAGCG 636
 595 GGAGCCAAAGCCAAACGAGCGGAGCTGAGAACGCTTACTAG 636
 637 GGAGCCAAAGCCAAACGAGCGGAGCTGAGAACGCTTACTAG 678

RESULT 6
 AX342535 2271 bp DNA linear PART 12-JAN-2002
 LOCUS
 DEFINITION Sequence 1 from Patent W00198491.
 ACCESSION AX342535
 VERSION AX342535.1 GI:18151964
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE
 1 Vickery,B.H.
 Osteolevin gene polymorphisms
 JOURNAL Patent: WO 0198491-A 1 27-DEC-2001;
 F. HOFMANN-LA ROCHE AG (CH) ; UNIVERSITAIRE INSTELLING ANTWERPEN
 (BE)
 FEATURES
 source Location/Qualifiers
 1..2271
 /organism="Homo sapiens"

BASE COUNT 573 a /db_xref="taxon:9606"
 623 c 599 g 476 t
 ORIGIN

Query Match 78.9%; Score 501.6; DB 6; Length 2271;
 Best Local Similarity 87.5%; Pred. No. 3.1e-76;
 Matches 562; Conservative 0; Mismatches 74; Indels 6; Gaps 1;

QY 1 ATGCAGCCCTCACTAGCCCGCTGCTCATCTGCTACTTGTGACGCTGCTCTGCTGCT 60
 DB 13 ATCAGCTCCCACTGCGCTCTGTGTCTGTCTGCTGTGTGTAACAACAGCTTCCGCTGTA 72
 QY 61 GTGAGGCGCCAGGGGTGGCAAGCTTCAAGATGATGCCACAGAGGTCAATCCAGGGCTT 120
 DB 73 GTGAGGCGCCAGGGGTGGCAAGCTTCAAGATGATGCCACAGAGGTCAATCCAGGGCTT 132
 QY 121 GGAGAGTACCCGAGGCTCTCTCC-----TGAGAACACAGACCATGTAACCGGGGGGAG 174
 DB 133 GGAGAGTACCCGAGGCTCTCTCC-----TGAGAACACAGACCATGTAACCGGGGGGAG 192
 QY 175 AATGAGGAGAGACCTCCCAACATCCCTATGACGCGCAAGATGTCTCAAGTACAGCTGC 234
 DB 193 AACGAGGAGGCGGCTCCCAACATCCCTATGACGCGCAAGATGTCTCAAGTACAGCTGC 252
 QY 235 CGGAGCTGCACTACACCGGCTTCTGACAGAGCGCCATGCGCGACGCGACCGGCTC 294
 DB 253 CGGAGCTGCACTACACCGGCTTCTGACAGAGCGCCATGCGCGACGCGACCGGCTC 312
 QY 295 ACCGAGTGTGTGTCTCCGCGCAGTGGCGCCGCGCGGCTGTGCGCCAAACGCTATCGG 354
 DB 313 ACCGAGTGTGTGTCTCCGCGCAGTGGCGCCGCGCGGCTGTGCGCCAAACGCTATCGG 372
 QY 355 CGGCTGAAGT 414
 DB 373 CGGCGCAAGT 432
 QY 415 GCGCAACGGGTGTGAGT 474
 DB 433 GCGCAACGGGTGTGAGT 492
 QY 475 CTGT 534
 DB 493 CTGT 552
 QY 535 GACTTGT 594
 DB 553 GACTTGT 612
 QY 595 GAGCGCAAGCCCAACGAGGCGGAGCTGTGAGAGAGGCGGAGGCGGCGGCGGCGG 636
 DB 613 AGCGCCAAAGCCCAACGAGGCGGAGCTGTGAGAGAGGCGGAGGCGGCGGCGGCGG 654

RESULT 7
 AF331844 2296 bp mRNA linear PRI 06-MAR-2001
 LOCUS AF331844
 DEFINITION Homo sapiens SOST (SOST) mRNA, complete cds.
 ACCESSION AF331844
 VERSION AF331844.1 GI:13236417
 KEYWORDS
 SOURCE
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 2296)
 Balcells, W., Ebeling, M., Patel, N., van Hul, E., Olsson, P.,
 Dierker, F. G., Stratakis, C., Lindpaintner, K., Vicky, B.,
 Foerzler, D. and Van Hul, W.
 Increased bone density in sclerosteosis is due to the deficiency of
 a novel secreted protein (SOST)
 Hum. Mol. Genet. 10 (5), 537-543 (2001)

MEDLINE 21096930
 PUBMED 1181578
 REFERENCE 2 (bases 1 to 2296)
 AUTHORS Balcells, W., Ebeling, M., Patel, N., Vicky, B., Foerzler, D. and Van
 Hul, W.
 TITLE Direct Submission
 JOURNAL Submitted (22-DEC-2000) Medical Genetics, University of Antwerp,
 Universiteitsplein 1, Antwerp B2610, Belgium
 FEATURES
 source
 1. 2296
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="17"
 /map="17q12-q21"
 1. 2296
 /gene="SOST"
 38. 679
 /gene="SOST"
 /note="Mutations may cause sclerosteosis"
 /codon_start=1
 /product="SOST"
 /protein_id="AAK16158.1"
 /db_xref="GI:13236418"
 /translation="MQLPLALCVLLVHTAFRVVSGQQAFAKNDATETIPLEGYR
 EPPLENNKTNNAENGSRPPHPPSTKDVSEKLEHFRVYDTPERSAKPTE
 LVSGDQGPALPLPNAIGRGKWRPSPDPRCTPDYRAORVOLLCPGSAPRARKVR
 LVASCKRRLRPFNOSLEKDPGTARPKGRPRPARSAKANQALENAY"

BASE COUNT 576 a 631 c 607 g 482 t
 ORIGIN

Query Match 78.9%; Score 501.6; DB 9; Length 2296;
 Best Local Similarity 87.5%; Pred. No. 3.1e-76;
 Matches 562; Conservative 0; Mismatches 74; Indels 6; Gaps 1;

QY 1 ATGCAGCCCTCACTAGCCCGCTGCTCATCTGCTACTTGTGACGCTGCTCTGCTGCT 60
 DB 38 ATCAGCTCCCACTGCGCTCTGTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 97
 QY 61 GTGAGGCGCCAGGGGTGGCAAGCTTCAAGATGATGCCACAGAGGTCAATCCAGGGCTT 120
 DB 98 GTGAGGCGCCAGGGGTGGCAAGCTTCAAGATGATGCCACAGAGGTCAATCCAGGGCTT 157
 QY 121 GGAGAGTACCCGAGGCTCTCTCC-----TGAGAACACAGACCATGTAACCGGGGGAG 174
 DB 158 GGAGAGTACCCGAGGCTCTCTCC-----TGAGAACACAGACCATGTAACCGGGGGAG 217
 QY 175 AATGAGGAGAGACCTCCCAACATCCCTATGACGCGCAAGATGTCTCAAGTACAGCTGC 234
 DB 218 AACGAGGAGGCGGCTCCCAACATCCCTATGACGCGCAAGATGTCTCAAGTACAGCTGC 277
 QY 235 CGGAGCTGCACTACACCGGCTTCTGACAGAGCGCCATGCGCGACGCGCAAGCGGCTC 294
 DB 278 CGGAGCTGCACTACACCGGCTTCTGACAGAGCGCCATGCGCGACGCGCAAGCGGCTC 337
 QY 295 ACCGAGT 354
 DB 338 ACCGAGT 397
 QY 355 GCGCTGAAGT 414
 DB 398 GCGCTGAAGT 457
 QY 415 GCGCAACGGGTGTGAGT 474
 DB 458 GCGCAACGGGTGTGAGT 517
 QY 475 CTGT 534
 DB 518 CTGT 577
 QY 595 GACTTGT 594
 DB 578 GACTTGT 637

QY 595 GGAGCCAAAGCCAGCGGAGCTGGAGAAAGCTCTAG 636
 Db 638 AGCGCCAAAGCCAGCGGAGCTGGAGAAAGCTCTAG 679
 RESULT 8
 LOCUS AF326739 2323 bp mRNA linear PRI 28-FEB-2001
 DEFINITION Homo sapiens sclerostin mRNA, complete cds.
 ACCESSION AF326739
 VERSION AF326739.1 GI:13161019
 KEYWORDS
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 2323)
 AUTHORS Brunkow, M.E., Gardner, J.C., Van Ness, J., Paepker, B.W.,
 Kovacevich, B.R., Prohl, S., Skonier, J.E., Zhao, L., Sabo, P.J.,
 Fu, Y.H., Altsch, R.S., Gillett, L., Colbert, T., Tacconi, P., Galas, D.,
 Hamersma, H., Beighton, P. and Mulligan, J.T.
 Bone dysplasia sclerosteosis results from loss of the SOST gene
 product, a novel cystine knot-containing protein
 Am. J. Hum. Genet. 68 (3), 577-589 (2001)
 JOURNAL MEDLINE 21090529
 PUBMED 11179006
 REFERENCE 2 (bases 1 to 2323)
 AUTHORS Brunkow, M.E., Gardner, J.C., Van Ness, J., Paepker, B.W.,
 Kovacevich, B.R., Prohl, S., Skonier, J.E., Zhao, L., Sabo, P.J.,
 Fu, Y.H., Altsch, R.S., Gillett, L., Colbert, T., Tacconi, P., Galas, D.,
 Hamersma, H., Beighton, P. and Mulligan, J.T.
 Direct Submission
 Submitted (07-DEC-2000) Genomics, Celltech Chiroscience Inc., 1631
 220th St. SE, Bothell, WA 98021, USA
 TITLE JOURNAL
 FEATURES
 source
 1. .2323
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="17"
 /map="17q21"
 1. .47
 48. .689
 /codon_start=1
 /product="sclerostin"
 /protein_id="AAK13454.1"
 /translation="MQLPLALCLVCLLVHTARVVEGQGMQAFKNDATETIIPELGEYP
 EPPLENNKTNRAENGSRPHHPETDVSYSCEHAFTRYVTDGPRRAKVT
 LVSGGCGPARLPNAIGRKMMRSPGPRCPIDRYRAORVOLLCPGGEAPRAKVR
 LVASCKCKRITRRHNSLKDPTFTEAARQKRRPRARSAKANQAELENNY"
 690. .2323
 BASE COUNT 590 a 635 c 614 g 484 t
 ORIGIN
 Query Match 78.9%; Score 501.6; DB 9; Length 2323;
 Best Local Similarity 87.5%; Pred. No. 3.1e-76;
 Matches 562; Conservative 0; Mismatches 74; Indels 6; Gaps 1;

Db 228 AACGAGGGCGGCTCCCAACCAACCTTTGAGACCAAAAGCTGTCCGATACAGCTGC 287
 QY 235 CGCAGCTGCATACACCCGTTCTTCTGACAGACGCCCATAGCGGCAAGCGGCTC 294
 Db 288 CGCAGCTGCATACACCCGTTCTTCTGACAGACGCCCATAGCGGCAAGCGGCTC 347
 QY 295 ACCGAGTTGGTGTCTCCGAGCCAGTGCAGCCCGCGGCTGCTGCCAAGCCATCGG 354
 Db 348 ACCGAGTTGGTGTCTCCGAGCCAGTGCAGCCCGCGGCTGCTGCCAAGCCATCGG 407
 QY 355 CGCAGTTGGTGTCTCCGAGCCAGTGCAGCCCGCGGCTGCTGCCAAGCCATCGG 414
 Db 408 CGCAGTTGGTGTCTCCGAGCCAGTGCAGCCCGCGGCTGCTGCCAAGCCATCGG 467
 QY 415 GCGAGCGGCGTGCAGCTGTGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 474
 Db 468 GCGAGCGGCGTGCAGCTGTGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 527
 QY 475 CTGCTGCGCTGCTGCGAGTGCAGCGCTTCAACCGCTTCCACACAGTGGAGCTCAAG 534
 Db 528 CTGCTGCGCTGCTGCGAGTGCAGCGCTTCAACCGCTTCCACACAGTGGAGCTCAAG 587
 QY 535 GACTTGGCGCGGAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGG 594
 Db 588 GACTTGGCGCGGAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGG 647
 QY 595 GGAGCCAAAGCCAGCGGAGCTGGAGAACGCTTACTAG 636
 Db 648 AGCGCCAAAGCCAGCGGAGCTGGAGAACGCTTACTAG 689
 RESULT 9
 LOCUS AX056687 2329 bp DNA linear PAT 17-JUN-2001
 DEFINITION Sequence 19 from Patent WO0075317.
 ACCESSION AX056687
 VERSION AX056687.1 GI:12309667
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 2329)
 AUTHORS Wood, W.I., Goddard, A., Gurney, A.L., Smith, V., Watanabe, C.K. and
 Wood, W.I.
 Compositions and methods for the treatment of tumor
 Patent: WO 0075317-A 19 14-DEC-2000;
 JOURNAL Genentech, Inc. (US)
 TITLE JOURNAL
 FEATURES
 source
 1. .2329
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 BASE COUNT 587 a 645 c 612 g 485 t
 ORIGIN
 Query Match 78.9%; Score 501.6; DB 6; Length 2329;
 Best Local Similarity 87.5%; Pred. No. 3.1e-76;
 Matches 562; Conservative 0; Mismatches 74; Indels 6; Gaps 1;

Db 242 AACGAGGCGGCGCTCCCAACCCCTTTTGAAGCAAAAGAGTGTCCAGTACGCTGC 301
 Qy 235 CGCGAGCTGCACTACACCGCTTCTCTGACAGACGCGCCATGCGCAGC3CCAAAGCGCTC 234
 Db 302 CGCGAGCTGCACTTACCCCGCTACGTGACCAATGAGCCGCTCCCGACAG3CCAAAGCGCTC 361
 Qy 235 ACCGAGTTGTGTGCTCTCCGCGCAGTGCAGCCCGCGCGCTGCTGCCAAAGCCATCGAG 354
 Db 362 ACCGAGCTGTGTGCTCTCCGCGCAGTGCAGCCCGCGCGCTGCTGCCAAAGCCATCGAG 421
 Qy 355 CGCGGAGAGTGTGTGCGCGCGCAACGCAATTTTCCGCTGATCCCGGATCCGCTACCGC 414
 Db 422 CGCGGAGAGTGTGTGCGCGCGCAATTTTCCGCTGATCCCGGATCCGCTACCGC 481
 Qy 415 GCGCAGCGGCTGCACTGCTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 474
 Db 482 GCGCAGCGGCTGCACTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 541
 Qy 475 CTGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 534
 Db 542 CTGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 601
 Qy 535 GACTTTCGCGCGCGAGACCG 594
 Db 602 GACTTTCGCGCGCGAGCG 661
 Qy 595 GGAGCCAAAGCCAAACGAGCGAGCTGAGAGAGCCCTACTAG 636
 Db 662 AGCCCAAGCCAAACGAGCGAGCTGAGAGAGCCCTACTAG 703
 RESULT 10
 AF326737 81806 bp DNA linear ROD 15-APR-2002
 LOCUS
 DEFINITION Mus musculus sclerostin (Sost) gene, complete cds; and Meox1 gene,
 partial sequence.
 ACCESSION AF326737
 VERSION AF326737.2 GI:20150447
 KEYWORDS
 SOURCE Mus musculus.
 ORGANISM Mus musculus.
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 81806)
 Brunkow, M.E., Gardner, J.C., Van Ness, J., Paepers, B.W.,
 Kovacevich, B.R., Prohl, S., Skonier, J.E., Zhao, L., Sabo, P.J.,
 Fu, Y.H., Alish, R.S., Gillet, L., Colbert, T., Tacconi, P., Galas, D.,
 Hamersma, H., Beighton, P. and Mulligan, J.T.
 Bone dysplasia sclerostosis results from loss of the SOST gene
 product, a novel cysteine knot-containing protein
 Am. J. Hum. Genet. 68 (3), 577-589 (2001)
 21090529
 JOURNAL MEDLINE
 PUBMED 11179006
 REFERENCE
 AUTHORS Staehling-Hampton, K., Prohl, S., Paepers, B.W., Zhao, L., Charnley, P.,
 Brown, A., Gardner, J.C., Galas, D., Schatzman, R.C., and Brunkow, M.E.
 A 52 kb deletion in the SOST - MEOX1 intergenic region on 17q12-q21
 is associated with van Buchem disease in the Dutch population
 Unpublished
 3 (bases 1 to 17423)
 Brunkow, M.E., Gardner, J.C., Van Ness, J., Paepers, B.W.,
 Kovacevich, B.R., Prohl, S., Skonier, J.E., Zhao, L., Sabo, P.J.,
 Fu, Y.H., Alish, R.S., Gillet, L., Colbert, T., Tacconi, P., Galas, D.,
 Hamersma, H., Beighton, P. and Mulligan, J.T.
 Direct Submission
 Submitted (07-DEC-2000) Genomics, Celltech Chiroscience Inc., 1631
 220th St. SE, Bothell, WA 98021, USA
 4 (bases 1 to 81806)
 Staehling-Hampton, K., Prohl, S., Paepers, B.W., Zhao, L., Charnley, P.,
 Brown, A., Gardner, J.C., Galas, D., Schatzman, R.C., Beighton, P.,
 Papapoulos, S., Hamersma, H. and Brunkow, M.E.
 Direct Submission
 TITLE

JOURNAL Submitted (02-AUG-2001) Genomics, Celltech R&D Inc., 1631 220th St
 SE, Bothell, WA 98021, USA
 REMARK Sequence update by submitter
 COMMENT On Apr 15, 2002 this sequence version replaced gi:13161013.
 FEATURES
 source
 1. 81806
 /organism="Mus musculus"
 /strain="129SV/7"
 /db_xref="taxon:10090"
 /chromosome="11"
 /map="between Sost and Meox1"
 <8036..21163
 /gene="Sost"
 join(<8036..8249,10742..21163)
 /gene="Sost"
 /product="sclerostin"
 join(8036..8249,10742..11163)
 /gene="Sost"
 /codon_start=1
 /product="sclerostin"
 /protein_id="AAK13452.1"
 /db_xref="GI:13161014"
 /translation="MPSLAPCLICLVHAFCAVEGGQGMQAFRDATVETPGLEIYP
 EPPENNQTNWNAENGRPPHPYDAKDVSEYSCHELYHTRFLTDPGRSAKPTVELV
 CSGCCGPARLPVIAIGRVKMRPNGDFPCIDRYARARVOLICGGAAPSRKRVRLV
 ASCKKRLTRFPHNOSLNDPQETARPOKGRPRGAGAKANQALENAI"
 81492..281806
 /gene="Meox1"
 81492..281806
 /gene="Meox1"
 /product="Meox1"
 BASE COUNT 20982 a 20323 c 20010 g 20491 t
 ORIGIN
 Query Match 67.1%; Score 426.6; DB 10; Length 81806;
 Best Local Similarity 95.8%; Pred. No. 1e-63;
 Matches 438; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
 Qy 180 AGGAGAGCTCCCGACCATCCTTATGACGCGCAAGATGTGCTGAGTACGTCGCGCGA 239
 Db 10707 AGGCTGACGCGCTTACGACATCTCTTCCGAGATGTGCTGAGTACGTCGCGCGA 10766
 Qy 240 GTTGACTACACCGCTTCTCTGACAGACGCGCCATGCGCAGCGCCAGCGGACCGA 299
 Db 10767 GCTGCACTACACCGCGCTTCTCTGACAGACGCGCCATGCGCAGCGGACCGA 10826
 Qy 300 GTTGATGTGCTCCGCGCAGTGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 359
 Db 10827 GTTGATGTGCTCCGCGCAGTGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 10886
 Qy 360 GAAAGTGTGCGCGCGCGCAACGAGCCGATTTCCGTCGATCCGCGATCGGTACCGCGCA 419
 Db 10887 GAAAGTGTGCGCGCGCGCAACGAGCCGATTTCCGTCGATCCGCGATCGGTACCGCGCA 10946
 Qy 420 GCGGATGACAGCTGCTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 479
 Db 10947 GCGGATGACAGCTGCTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 11006
 Qy 480 GCGCTGTGTAATGTGCAAGCGCTTACCGCGCTTCCAAACAGTGTGAGCTCAAGACTT 539
 Db 11007 GCGCTGTGTAATGTGCAAGCGCTTACCGCGCTTCCAAACAGTGTGAGCTCAAGACTT 11066
 Qy 540 CGGCGCGAGACCGCGCGCGCGCGAGAGGTGTGCAAGCGCGCGCGCGCGCGCGCGCGCG 599
 Db 11067 CGGCGCGAGACCGCGCGCGCGCGAGAGGTGTGCAAGCGCGCGCGCGCGCGCGCGCG 11126
 Qy 600 CAAAGCCAAACGAGCGGAGCTGAGAGACGCTTACTAG 636
 Db 11127 CAAAGCCAAACGAGCGGAGCTGAGAGACGCTTACTAG 11163
 RESULT 11
 AC068782_2

WPCOMMENT

Sequence split into 5 fragments LOCUS AC068782 Accession AC068782

Fragment Name Begin End

AC068782_0 1 110000

AC068782_1 100001 210000

AC068782_2 200001 310000

AC068782_3 300001 410000

AC068782_4 400001 411130

Continuation (3 of 5) of AC068782 from base 200001 (AC068782 Mus musculus chromosome 11

Query Match 67.1%; Score 426.6; DB 2; Length 110000;

Best Local Similarity 95.8%; Pred. No. 9.7e-64;

Matches 438; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 180 AGGACAGCTCTCCACCATCTCTATGACCGCAAGATGTGTCCGAGTACAGCTGCCGCGA 239

DB 56567 AGGCTCAGACCCCTCAGCATCCCTCTCCGAGATGTGTCCGAGTACAGCTGCCGCGA 56626

QY 240 GCTGCACTACACCGCTTCTGACAGAGCGCCATGCGGAGCCGAGCCGATCACCGA 299

DB 56627 GCTGCACTACACCGCTTCTGACAGAGCGCCATGCGGAGCCGAGCCGATCACCGA 56686

QY 300 GTTGTGTGTCTCCGCGCAAGTGCAGCGCCCGCGCGCTGTCTGCCCAACGCCATCGGCGCGT 359

DB 56687 GTTGTGTGTCTCCGCGCAAGTGCAGCGCCCGCGCGCTGTCTGCCCAACGCCATCGGCGCGT 56746

QY 360 GAAGTGTGTGTCTCCGCGCAAGTGCAGCGCCCGCGCGCTGTCTGCCCAACGCCATCGGCGCGT 419

DB 56747 GAAGTGTGTGTCTCCGCGCAAGTGCAGCGCCCGCGCGCTGTCTGCCCAACGCCATCGGCGCGT 56806

QY 420 GCGGAGTGTGTCTGTGTGCG 479

DB 56807 GCGGAGTGTGTCTGTGTGCG 56866

QY 480 GGCCTGT 539

DB 56867 GGCCTGT 56926

QY 540 CGGCGCGGAGACCG 599

DB 56927 CGGCGCGGAGACCG 56986

QY 600 CAAAGCCAAACGAGCGCGAGCTGTGAGAAAGCGCTACTAG 636

DB 56987 CAAAGCCAAACGAGCGCGAGCTGTGAGAAAGCGCTACTAG 57023

RESULT 12 AC012296 205277 bp DNA linear HTG 16-AUG-2002

LOCUS AC012296 Mus musculus chromosome 11 clone RP23-346P7 map 11, *** SEQUENCING

DEFINITION IN PROGRESS ***, 10 unordered pieces.

AC012296 205277 bp DNA linear HTG 16-AUG-2002

AC012296.10 GI:22267737

HTG: HTGS_PHASE1; HTGS_FULLTOP; HTGS_ACTIVEFIN.

KEYWORDS house mouse.

SOURCE Mus musculus

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 205277)

Birren,B., Nusbaum,C. and Lander,E.

Mus musculus chromosome 11, clone RP23-346P7

Unpublished

2 (bases 1 to 205277)

Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M., Baldwin,J., Barna,N., Beckerly,R., Boguslavsky,L., Bouhassira,B., Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A., Cooke,P., Dearlano,K., Dewar,K., Domino,M., Donelan,J., Doyle,M., Ferreira,P., Fitzhugh,W., Forrest,C., Funke,R., Gage,D., Galagan,J., Gardyna,S., Grant,G., Hages,B., Heathford,A., Horton,L., Howland,J.C., Johnson,R., Jones,C., Kamm,L., Karatas,A., Klein,J., Lehotzky,J., Lien,C., Locke,K., Macdonald,P., Marquis,N., McSwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrum,J.,

TITLE

Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P., Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P., Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J., Testaye,S., Tittell,A., Vasilev,H., Vo,A., Wheeler,J., Wu,X., Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.

REFERENCE

Submitted (22-OCT-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 205277)

AUTHORS

Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Bouhassira,B., Camarata,J., Chang,J., Chazaro,B., Choepel,T., Collymore,A., Cook,A., Cooke,P., Dearlano,K., Dewar,K., Diaz,J.S., Dodge,S., Fero,S., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hages,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Maclean,C., Macdonald,P., Major,J., Matthews,C., McCarthy,M., Meldrum,J., Menais,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunhahng,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schupbach,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J., Testaye,S., Theodore,J., Topham,K., Travers,M., Vasilev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

COMMENT

Submitted (16-AUG-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

On Aug 16, 2002 this sequence version replaced gi:20177756.

All repeats were identified using RepeatMasker:

Smith, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

JOURNAL

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: MIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu

Project Information

Center project name: L3205

Center clone name: 346_P_7

***** NOTE: This is a 'working draft' sequence. It currently consists of 10 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 92178: contig of 92178 bp in length

* 92179 92278: gap of 100 bp

* 92279 128704: contig of 36426 bp in length

* 128705 128804: gap of 100 bp

* 128805 140272: contig of 11468 bp in length

* 140273 140372: gap of 100 bp

* 140373 147001: contig of 6629 bp in length

* 147002 147101: gap of 100 bp

* 147102 170176: contig of 23075 bp in length

* 170177 170276: gap of 100 bp

* 170277 173609: contig of 3333 bp in length

* 173610 173709: gap of 100 bp

* 173710 175722: contig of 2013 bp in length

* 175723 175822: gap of 100 bp

* 175823 181949: contig of 6127 bp in length

* 181950 182049: gap of 100 bp

* 182050 191157: contig of 9108 bp in length

* 191158 191257: gap of 100 bp

* 191258 205277: contig of 14020 bp in length.

FEATURES

1..205277

/organism="Mus musculus"

source


```
/rpt_family="PB1"
repeat_region 9190..9358
/rpc_family="MER2"
repeat_region 9359..9707
/rpc_family="ORR1D"
repeat_region 9708..9801
/rpc_family="MER2"
repeat_region 9802..9831
/rpc_family="(TTTG)n"
repeat_region 9832..9886
/rpc_family="MER2"
repeat_region 10219..10239
/rpc_family="(TG)n"
repeat_region 12931..13108
/rpc_family="L1MC4"
repeat_region 13166..13231
/rpc_family="MIR"
repeat_region 14252..14426
/rpc_family="MIR"
repeat_region 14911..15037
/rpc_family="L2"
repeat_region 15543..15794
/rpc_family="B4"
repeat_region 15849..15890
/rpc_family="(CA)n"
repeat_region 15971..16099
/rpc_family="L2"
repeat_region 16336..16404
/rpc_family="BC1_MM"
repeat_region 16419..16550
/rpc_family="B1-F"
repeat_region 17412..17518
/rpc_family="B1-F"
repeat_region 17538..17710
/rpc_family="B3A"
repeat_region 18267..18397
/rpc_family="B3A"
repeat_region 18508..18679
/rpc_family="B3"
repeat_region 18720..18932
/rpc_family="B3"
repeat_region 18939..18982
/rpc_family="(TTAG)n"
repeat_region 18989..19122
/rpc_family="B1_MM"
repeat_region 20446..20476
/rpc_family="(TG)n"
repeat_region 20593..20889
/rpc_family="B4A"
repeat_region 20964..21091
/rpc_family="PB1D10"
repeat_region 22076..22187
/standard_name="163518"
repeat_region 22629..22743
/rpc_family="MIR"
repeat_region 22748..22899
/rpc_family="B4A"
repeat_region 23282..23305
/rpc_family="(A)n"
repeat_region 23322..23355
/rpc_family="B2"
repeat_region 23381..23525
/rpc_family="B2"
repeat_region 24185..24255
/rpc_family="(TG)n"
repeat_region 24921..25202
/rpc_family="Lx6"
repeat_region 25271..25346
/rpc_family="B3"
repeat_region 25352..25444
/rpc_family="PB1D10"
repeat_region 26173..26320
/rpc_family="L1MD"
```

```
repeat_region 26434..26601
/rpc_family="RSINE1"
repeat_region 26720..26801
/rpc_family="B4A"
repeat_region 26802..26906
/rpc_family="PB1D7"
repeat_region 27067..27454
/rpc_family="L1_MM"
repeat_region 27555..27575
/rpc_family="(A)n"
repeat_region 27956..28217
/rpc_family="B4"
repeat_region 28275..28459
/rpc_family="L1M4"
repeat_region 28460..28615
/rpc_family="B3A"
repeat_region 28937..29077
/rpc_family="B2"
repeat_region 29241..29376
/rpc_family="L1M4"
repeat_region 29323..29410
/rpc_family="B4A"
repeat_region 29438..29691
/rpc_family="PB1D10"
repeat_region 29696..29823
/rpc_family="(CA)n"

Query Match 67.1%; Score 426.6; DB 10; Length 208135;
Best Local Similarity 95.8%; Pred. No. 8.7e-64;
Matches 438; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 180 AGCGAGACCTCCCAACCAATCCCTATGAGCCCAAGATGTGCGAGTACAGTCCCGCGA 239
DB 57814 AGCTGACAGCCCTCCCAACCAATCCCTATGAGCCCAAGATGTGCGAGTACAGTCCCGCGA 57873
QY 240 GGTGACATACACCGCTTCTGACAGACGCGCCCAATGCGCGACCGCAAGCGGTACCGA 299
DB 57874 GTTGACATACACCGCTTCTGACAGACGCGCCCAATGCGCGACCGCAAGCGGTACCGA 57933
QY 300 GTTGTGTGCTCCCGGCAAGTGGGCGCGCGGCTGCTGCCAAGCGCATCGGCGCGT 359
DB 57934 GTTGTGTGCTCCCGGCAAGTGGGCGCGCGGCTGCTGCCAAGCGCATCGGCGCGT 57993
QY 360 GAAGTGTGCGCGCGCGGCAAGTGGGCGCGGCTGCTGCCAAGCGCATCGGCGCGA 419
DB 57994 GAAGTGTGCGCGCGCGGCAAGTGGGCGCGGCTGCTGCCAAGCGCATCGGCGCGA 58053
QY 420 GCGGTGACAGTGTGCTGCGCGCGGCGCGCGCGCTGCGCGCAAGTGTGCTGCT 479
DB 58054 GCGGTGACAGTGTGCTGCGCGCGGCGCGCGCGCTGCGCGCAAGTGTGCTGCT 58113
QY 480 GCGCTGTGCAAGTCAAGCGGCTCAACCGCTTCCCAACCAAGTGTGCAAGGACTT 539
DB 58114 GCGCTGTGCAAGTCAAGCGGCTCAACCGCTTCCCAACCAAGTGTGCAAGGACTT 58173
QY 540 CCGGCGCGGAGACCGCGCGCGGCGGCAAGGTCGCAAGCGCGCGCGCGCGGAGC 599
DB 58174 CCGGCGCGGAGACCGCGCGCGGCGGCAAGGTCGCAAGCGCGCGCGCGGAGC 58233
QY 600 CAAAGCAACGAGCGGAGCTGAGAAAGCTTACTAG 636
DB 58234 CAAAGCAACGAGCGGAGCTGAGAAAGCTTACTAG 58270

RESULT 14
AL591145 198508 bp DNA linear ROD 22-AUG-2002
LOCUS AL591145
DEFINITION Mouse DNA sequence from clone Rp23-398T7 on chromosome 11, complete
ACCESSION AL591145
VERSION AL591145.24 GI:22474402
```


KEYWORDS HTG.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE 1 (bases 1 to 198508)
JOURNAL Direct Submission
COMMENT Submitted (22-AUG-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Aug 23, 2002 this sequence version replaced gi:22003099.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk

During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em: EMBL; Sw:
SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-398F7 is
from the RP23-Mouse PAC Library
constructed by the group of Pieter de Jong.
For further details see http://www.chori.org/bacpac/home.htm
VECTOR: pBACe3.6.
Location/Qualifiers
1..198508
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="11"
/clone="RP23-398F7"
/clone_1b="RP23-23"
BASE COUNT 48902 a 47818 c 50456 g 51332 t
ORIGIN
Query Match 66.8%; Score 425; DB 10; Length 198508;
Best Local Similarity 95.6%; Pred. No. 1.6e-63;
Matches 437; Conservative 0; Mismatches 20; Indels 0; Gaps 0;
QY 180 AGGAGACCTTCCCAACCATCCCTATGACGCCAAAGATGTGTCCGAGTACAGCTGCCGCGA 239
DB 187711 AGGCTGACGCCCTTCACGACATCCCTCTCTCCGAGATGTGTCCGAGTACAGCTGCCGCGA 187770
QY 240 GCTGCACTACACCGGCTTCCGACAGAGCGGCGCATGCGAGGCGCAAGCGGCTACACCGA 239
DB 187771 GCTGCACTACACCGGCTTCCGACAGAGCGGCGCATGCGAGGCGCAAGCGGCTACACCGA 187830
QY 300 GTTGTGTGTCTCCGCGGACGATGCGGCGCGGCTGTCTCCCAACGCGCATCGGCGCGCT 359
DB 187831 GTTGTGTGTCTCCGCGGACGATGCGGCGCGGCTGTCTCCCAACGCGCATCGGCGCGCT 187890
QY 360 GAAGTGTGTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 419
DB 187891 GAAGTGTGTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 187950
QY 420 GCGGCTCAGCTGTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 479
DB 187951 GCGGCTCAGCTGTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 188010

QY 480 GGCTCTGTGCAAGTGTGCAAGCGGCTTCCGACAAACGAGTGTGAGCTTCAAGACTT 539
DB 188011 GGCTCTGTGCAAGTGTGCAAGCGGCTTCCGACAAACGAGTGTGAGCTTCAAGACTT 188070
QY 540 CGGCGCGGAGACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 599
DB 188071 CGGCGCGGAGACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 188130
QY 600 CAAAGCAACACGAGCGGAGCTGAGAGAGCTTACTAG 636
DB 188131 CAAAGCAACACGAGCGGAGCTGAGAGAGCTTACTAG 188167
RESULT 15
AF326738 532 bp mRNA linear MAM 28-FEB-2001
LOCUS AF326738
DEFINITION Bos taurus sclerostin mRNA, partial cds.
ACCESSION AF326738
VERSION AF326738.1 GI:13161016
KEYWORDS
SOURCE Bos taurus.
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
REFERENCE
AUTHORS Kovacevich, B.R., Proli, S., Skonier, J.E., Zhao, L., Sabo, P.J.,
Fu, Y.H., Alisch, R.S., Gillett, L., Colbert, T., Taccott, P., Galas, D.,
Hamerma, H., Beighton, P. and Mulligan, J.T.
Bone dysplasia sclerosteosis results from loss of the SOST gene
product, a novel cycline knot-containing protein
Am. J. Hum. Genet. 68 (3), 577-589 (2001)
JOURNAL
MEDLINE 21090529
PUBMED 11179006
REFERENCE 2 (bases 1 to 532)
AUTHORS Brunkow, M.E., Gardner, J.C., Van Ness, J., Paepert, B.W.,
Kovacevich, B.R., Proli, S., Skonier, J.E., Zhao, L., Sabo, P.J.,
Fu, Y.H., Alisch, R.S., Gillett, L., Colbert, T., Taccott, P., Galas, D.,
Hamerma, H., Beighton, P. and Mulligan, J.T.
Direct Submission
Submitted (07-DEC-2000) Genomics, Celtech Chiroscience Inc., 1631
220th St. SE, Bothell, WA 98021, USA
Location/Qualifiers
1..532
/organism="Bos taurus"
/db_xref="taxon:9913"
<1..>532
/codon_start=3
/product="sclerostin"
/protein_id="AAK13453.1"
/db_xref="GI:13161017"
/translation="NDATETITPELGEPPEPELNNKTNRAENGRRPHHPETKTA
SEYSCREHLEFRTVYDGPGRSAKPVTVDSGQGPRLPLFNALGRKMWPPSPDR
CIPDRAORVOLLCPGGAAPRAKRVLVASCCKRLTRFNOSLKDQFPEARPT
GRKLRPRRGTKASRA"
BASE COUNT 95 a 199 c 176 g 62 t
ORIGIN
Query Match 65.4%; Score 415.8; DB 4; Length 532;
Best Local Similarity 87.8%; Pred. No. 1.7e-61;
Matches 466; Conservative 0; Mismatches 62; Indels 3; Gaps 1;
QY 90 GAATATGCAAGAGATGATCCCGGCTTGGAGATGACCCGAGCTTCTCTGA--- 146
DB 2 GAATATGCAAGAGATGATCCCGGCTTGGAGATGACCCGAGCTTCTCTGA--- 61
QY 147 GAACAACGACATGAACCGGCGGAGAAATGAGAGCGAGACTCCCAACATCCCTATTA 206
DB 62 GAACAACGACATGAACCGGCGGAGAAATGAGAGCGAGACTCCCAACATCCCTATTA 121
QY 207 GCGCAAGATGTGTCCGAGTACAGCTGCCGCGAGCTGACATACCGCTTCTGACAGA 266

```

Db 122 GACCAAGAGCCCTCCGAGTACAGCTGCAGGAGCTGCACTTCAACCCGCTACGTGACCGA 181
QY 267 CGGCCCATGCGCGAGCGCCCAAGCCGGTACCCGAGTTGTGTCTCCGACCAAGTGGGCC 326
Db 182 TGGGCGGTGCGCGAGCGCCCAAGCCGGTACCCGAGTTGTGTCTCCGACCAAGTGGGCC 241
QY 327 CGCGGAGCTGTGCTGCAAGCGCATCGGGCGGGTGAAGTGTGGGCGCCGAAACGGACCGGA 386
Db 242 GCGCGGCTGTGCTGCAAGCGCATCGGGCGGGTGAAGTGTGGGCGCCGAAACGGACCGGA 301
QY 387 TTTCGCTGCATCCCGAGTGCCTACCGCGCGAGCGGGTGCAGCTGTGTGCCCGGGGG 446
Db 302 CTTCGCTGCATCCCGAGTGCCTACCGCGCGAGCGGGTGCAGCTGTGTGTCTCTGGCGG 361
QY 447 CGCGGCGCGCGGCTCGCGCAAGTGCCTGTGTGGCTCTCGTGCAGTGCAGCGCTTAC 506
Db 362 CGCGGCGCGCGGCTCGCGCAAGTGCCTGTGTGGCTCTCGTGCAGTGCAGCGCTTAC 421
QY 507 CGCTTCCACAACGAGTGCAGTCTCAAGGACTTTCGGGCGGGAGACCGCAACCAAGGCGGA 566
Db 422 TCGCTTCCACAACGAGTGCAGTCTCAAGGACTTTCGGGCGGGAGACCGCAACCAAGGCGGA 481
QY 567 GGGTGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 617
Db 482 GGGCGGAGGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 532

```

Search completed: March 29, 2003, 00:43:58
 Job time : 2293.01 secs

Db	240	GSCCCCGGCGAGNTGTGGCCAACGCATCTGGAGCGGTGAAGTGGTGAGGCCCAACGSA	181
Oy	424	CCCCAATTCCGTCGATCCCCGACCGCTAACCGCGCACCGGTGACAGTTGCTGTGCC	483
Db	180	CCGACTTCCTCGGTGATCCCGATCGTTAACCGCGCACCGGATGAGCTGTGTCGTGCC	121
Oy	484	GGTGTGAGAGCCCGCGCGCGCAGAAGTGGCGCTGTGTGGCTTCGTGCAAGTGAAGCGC	543
Oy	544	CTCACCGGTTCCACAACCATGTGGAGCTCAAGACTTCGAGACCGAGCCCGCTGGCGC	603
Db	60	CTCACCGGTTCCACAACCATGTGGAGCTCAAGACTTCGAGCTGAGACCGCTGGCGC	1
RESULT 7	A11131/c		
LOCUS	A11131.1	315 bp	mRNA linear EST 11-FEB-1999
DEFINITION	UI-R-C2P-nr-f-06-0-UI s1 UI-R-C2P Rattus norvegicus cDNA clone		
ACCESSION	U1-R-C2P-nr-f-06-0-UI 3'		mRNA sequence.
VERSION	A11131.1		
KEYWORDS	A11131.1 GI:3513080		
SOURCE	EST.		
ORGANISM	Norway rat. Rattus norvegicus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus. 1 (bases 1 to 315) Bonaldo,M.F., Lennon,G. and Soares,M.B. Normalization and subtraction: two approaches to facilitate gene discovery Genome Res. 6 (9), 791-806 (1996)		
REFERENCE	JOURNAL MEDLINE COMMENT		
AUTHORS	Contact: Soares, MB Program for Rat Gene Discovery and Mapping University of Iowa 451 Eckstein Medical Research Building Iowa City, IA 52242, USA Tel.: 319 335 8250 Fax: 319 335 9565 Email: msoares@blue.weeg.uiowa.edu		
TITLE	Oligo-dT track not found, Not I site shown in beginning of sequence is likely internal to the message. cDNA Library Preparation: M. Fatima Bonaldo, Ph.D. Clone distribution: clones will be available through Research Genetics The following repetitive elements were found in this cDNA sequence: 97-125, >Gc_richlow_complexity Seq primer: M13 Forward.		
FEATURES	Location/Qualifiers		
SOURCE	1..315 /organism="Rattus norvegicus" /strain="Sprague-Dawley" /db_xref="taxon:10116" /clone="UI-R-C2P-nr-f-06-0-UI" /clone_lib="UI-R-C2P" /dev_stage="adult" /lab_host="DH10B (Life Technologies)" /note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-C2P library is a subtracted library derived from the UI-R-C1 library, which is a subtracted library derived from the UI-R-C0 library. The UI-R-C0 library consisted of a mixture of individually tagged normalized libraries constructed from rat placenta, adult lung, brain, liver, kidney, heart, spleen, ovary, muscle, 8, 12 and 18-day embryo. The tag is a string of 3-5 nucleotides present between the Not I site and the oligo-dT track which allows identification of the library of origin of a clone within the mixture. The subtracted library (UI-R-C2P) was constructed as follows: PCR amplified cDNA inserts from UI-R-C1 clones from which 3' ESTs had been derived was used as a driver in a hybridization with the UI-R-C1 library in the form of single-stranded circles. The remaining single-stranded circles (subtracted library) was		

BASE COUNT	46 a	110 c	116 g	43 t	
ORIGIN	purified by hydroxyapatite column chromatography, converted to double-stranded circles and electroporated into DH10B bacteria (Life Technologies) to generate the <i>U1-R-32p</i> library. This procedure has been previously described (Bonaldi, Lennon and Soares, Genome Research 6: 791-806, 1996)"				
Query Match	33.7%	Score 255.8	DB 9	Length 315	
Best Local Similarity	88.3%	Pred. No. 1.2e-42			
Matches 278	Conservative 0	Mismatches 37	Indels 0	Gaps 0	
QY	289	TTACCCGCTACGTGACCGATGAGCGCCGTGCGCAGCGCCAAAGCCGGTACCGAGCTGTG 348			
Db	315	TACACCCGCTTCGTGTAACCGACGCGCCGTGCGCAGTGCACCAAGCCGATCACCGAGTTGGTG 256			
QY	349	TGCTCCGGGCGAGTGGGCGCCCGCGGCGCTGTGCTCCCAACGCATATGCGCGCGCAATGG 408			
Db	255	TGCTCGGGCCAGTGGGCGCCCGCGGCGCTGTGCTCCCAACGCATATGCGCGCGCTGAAATGG 196			
QY	409	TGCGCACTAGTGGGCGCGGACCTTCCGCTGCATATCCCGACCGCTACCGCGCGAGCGCTG 468			
Db	195	TGGCGCCCGGAACGACCCGACCTTCCGCTGCATATCCGATATCGTACCGCGCGAGCGGGTG 136			
QY	469	CAGCTGCTGTGTCCCGATGTGAAGCGCGCGCGCGCGCCAAAGTGGCGGCTGTGAGCTTGG 528			
Db	135	CAGCTGCTGTATCCCGCGCGCGCGCGCGCGCGCTGCAGTATGATGATGAGTGCCTCG 76			
QY	529	TGCAAGTGAAGCGGCTTACCCGCTTCCCAACAGATGGAGCTTCAAGACTTCCGAGAC 588			
Db	75	TGCAAGTGAAGCGGCTTACCCGCTTCAACCAAGTGGAGCTTCAAGGACTTCCGAGACT 16			
QY	589	GAGGCGGCTTCGCGCG 603			
Db	15	GAGACCGTGGCCCGC 1			
RESULT 8	535 bp DNA linear GSS 17-OCT-1998				
LOCUS	AO171546/c				
DEFINITION	HS_3088_B1_H06_MF CIT Approved Human Genomic Sperm library D Homo sapiens genomic clone Plate=3088 Col=11 Row=P, DNA sequence.				
ACCESSION	AO171546				
VERSION	AO171546.1 GI:3568913				
KEYWORDS	GSS.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS	1 (bases 1 to 535) Mahairas,G.G., Wallace,J.C., Smatch,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.				
TITLE	Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome				
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)				
COMMENT	99380589 Contact: Mahairas GG, Wallace JC, Hood L High Throughput Sequencing Center University of Washington 401 Queen Anne Avenue North, Seattle, WA 98109, USA Tel: (206) 616-3618 Fax: (206) 616-3887 Email: jwallace@u.washington.edu Sequence Tagged Connector Plate: 3088 row: P column: 11 Class: BAC ends High quality sequence stop: 535. Location/Qualifiers 1. 535 /organism="Homo sapiens" /db_xref="taxon:9606"				
FEATURES	source				

/clone="Plate=3088 Col=11 Row=P"
 /clone.lib="CIT Approved Human Genomic Sperm Library D"
 /sex="Male"
 /note="Organ: sperm; Vector: pBelobAC11; BAC Clones in
 E-Coli DH10B"
 BASE COUNT 98 a 164 c 139 g 132 t 2 others
 ORIGIN

Query Match 33.1%; Score 251.2; DB 17; Length 535;
 Best Local Similarity 98.8%; Pred. No. 1e-41;
 Matches 253; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

1 TACTGAGAGTGGCGTCCCTCTCTGCTGCTGACATGAGCTCCCATGCGCTGTGT 60
 304 TACTGAGAGTGGCGTCCCTCTCTGCTGCTGACATGAGCTCCCATGCGCTGTGT 245
 61 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
 244 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 185
 121 TTCAAGAAATATGCGACGAAATCATCCCGAGCTCGGAGATGACCCCGAGCTCCACCG 180
 184 TTCAAGAAATATGCGACGAAATCATCCCGAGCTCGGAGATGACCCCGAGCTCCACCG 125
 181 GAGCTGAGAAACAACAAGCATGAAACCGGCGGAGAAACGAGGCGGCTCCCAACAC 240
 124 GAGCTGAGAAACAACAAGCATGAAACCGGCGGAGAAACGAGGCGGCTCCCAACAC 65
 241 CCCTTTGAGACCAAG 256
 64 ACCTTTGAGACCAAG 49

RESULT 9
 BE111224/c 291 bp mRNA linear EST 13-JUN-2000
 LOCUS BE111224
 DEFINITION UI-R-BJ1-auw-b-04-0-UI.s1 UI-R-BJ1 Rattus norvegicus cDNA clone
 BE111224
 VERSION UI-R-BJ1-auw-b-04-0-UI 3', mRNA sequence.
 BE111224.1 GI:8503329

KEYWORDS
 EST.
 SOURCE Norway rat.
 ORGANISM Rattus norvegicus
 Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE
 1 (bases 1 to 291)
 Bonaldo,M.F., Lennon,G. and Soares,M.B.
 Normalization and subtraction: two approaches to facilitate gene
 discovery

JOURNAL
 MEDLINE
 COMMENT Genome Res. 6 (9), 791-806 (1996)
 97044477

Contact: Soares, MB
 Program for Rat Gene Discovery and Mapping
 University of Iowa
 451 Eckstein Medical Research Building Iowa City, IA 52242, USA
 Tel.: 319 335 8250
 Fax: 319 335 9565
 Email: msoares@blue.weeg.uiowa.edu
 Oligo-dT track not found, Not 1 site shown in beginning of sequence
 is likely internal to the message. cDNA library preparation: M.B.
 Soares Lab Clone distribution: clones will be available through
 Research Genetics (www.resgen.com) The following repetitive
 elements were found in this cDNA sequence: 97-125,
 >CC rich low complexity
 Seq primer: M13 Forward
 POLYA=No.

FEATURES
 source
 1. .291
 /organism="Rattus norvegicus"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /clone="UI-R-BJ1-auw-b-04-0-UI"

/clone.lib="UI-R-BJ1"
 /lab host="DH10B (Life Technologies)"
 /note="Vector: p773D-Pac (Pharmacia) with a modified
 polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-BJ1
 library is a subtracted library derived from the following
 tissues: atrium at 16.5 dpc, ventricle at 16.5 dpc, AV
 canal at 16.5 dpc, atrium at 15 dpc, ventricle at 15 dpc,
 AV canal at 15 dpc, ventricle at 13 dpc, and adult heart.
 For a detailed description of the library from which this
 clone was derived, please visit our web site at
 ratseq.eng.uiowa.edu. The subtraction has been previously
 described in (Bonaldo, Lennon and Soares, Genome Research
 6:791-806, 1996)
 TAG SEQ=None found"

BASE COUNT 41 a 101 c 109 g 40 t
 ORIGIN

Query Match 32.0%; Score 243; DB 10; Length 291;
 Best Local Similarity 89.7%; Pred. No. 4.8e-40;
 Matches 261; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

313 CGGTCCGCGAGCGCCCAAGCGGTGACGAGCTGTGTGCTCCGCGCATGCGCGCGCG 372
 291 CGGTCCGCGAGCGCCCAAGCGGTGACGAGCTGTGTGCTCCGCGCATGCGCGCGCG 232
 373 CGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 432
 231 CGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 172
 433 CGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 492
 171 CGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 112
 493 GCG 552
 111 GCG 52
 553 TTCCACAACAGTCGAGCTCAAGACTTTCGAGACCGAGCGCGCTCGCGCG 603
 51 TTCCACAACAGTCGAGCTCAAGACTTTCGAGACCGAGCGCGCGCGCGCG 1

RESULT 10
 BF523030 254 bp mRNA linear EST 11-DEC-2000
 LOCUS BF523030
 DEFINITION UI-R-C2P-rk-a-03-0-UI.r1 UI-R-C2P Rattus norvegicus cDNA clone
 UI-R-C2P-rk-a-03-0-UI 5', mRNA sequence.
 BF523030
 ACCESSION BF523030.1 GI:11631045
 VERSION BF523030.1
 KEYWORDS EST.

SOURCE
 ORGANISM Norway rat.
 Rattus norvegicus
 Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE
 1 (bases 1 to 254)
 Bonaldo,M.F., Lennon,G. and Soares,M.B.
 Normalization and subtraction: two approaches to facilitate gene
 discovery

JOURNAL
 MEDLINE
 COMMENT Genome Res. 6 (9), 791-806 (1996)
 97044477

Contact: Soares, MB
 Program for Rat Gene Discovery and Mapping
 University of Iowa
 451 Eckstein Medical Research Building Iowa City, IA 52242, USA
 Tel.: 319 335 8250
 Fax: 319 335 9565
 Email: msoares@blue.weeg.uiowa.edu
 cDNA library preparation: M.B. Soares Lab Clone distribution:
 clones will be available through Research Genetics (www.resgen.com)
 This clone is also available through the I.M.A.G.E. Consortium at
 LIND (info@image.llnl.gov). IMAGE ID= 1793134
 Seq primer: M13 Forward.

Db	596	TTTGAAGGTGTGCGCAGCCAGCCCGCCGACGACGACGAGGAGGAGGAGCAGC	655
Qy	637	AGCGCCA	643
Db	656	ATATCCA	662
RESULT 12			
LOCUS	AK002396	1612 bp	mRNA
DEFINITION	Mus musculus adult male kidney cDNA, RIKEN full-length enriched library, clone:0610009G22:homolog to HYPOTHETICAL 15.3 KDA PROTEIN (FRAGMENT), full insert sequence.		
ACCESSION	AK002396		
VERSION	AK002396.1	GI:12832347	
KEYWORDS	HTC; CAP trapper.		
SOURCE	Mus musculus (strain:C57BL/6J) adult male kidney cDNA to mRNA, clone 11b:RIKEN full-length enriched mouse cDNA library clone:0610009G22.		
ORGANISM	Mus musculus		
REFERENCE	Mammalia; Euteleostomi; Chordata; Craniata; Vertebrata; Euteleostomi; Eumetazoa; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	Carninci, P. and Hayashizaki, Y.		
TITLE	1		
JOURNAL	High-efficiency full-length cDNA cloning		
MEDLINE	Meth. Enzymol. 303, 19-44 (1999)		
PUBMED	99279253		
REFERENCE	10349636		
AUTHORS	2		
TITLE	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.		
JOURNAL	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes		
MEDLINE	Genome Res. 10 (10), 1617-1630 (2000)		
PUBMED	20499374		
REFERENCE	11042159		
AUTHORS	3		
TITLE	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Kitenaka, T., Tashiro, H., Itoh, M., Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kaishiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Matsubiki, K., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y.		
JOURNAL	RIKEN integrated sequence analysis (RISA) system-384-Format		
MEDLINE	Genome Res. 10 (11), 1757-1771 (2000)		
PUBMED	20530913		
REFERENCE	11076861		
AUTHORS	4		
TITLE	Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Aikawa, T., Hara, A., Fukunishi, Y., Kono, H., Adachi, J., Fukuda, S., Aizawa, K., Iwata, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Pleischmann, M., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuweli, P., Lewis, S., Matsuo, Y., Nikaio, I., Pesole, G., Quackenbush, J., Schriml, L. M., Stabli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Matsuura, M., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyooka, K., Wang, K. H., Wetz, C., Whitaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kontseki, S. and Hayashizaki, Y.		
JOURNAL	Functional annotation of a full-length mouse cDNA collection		
MEDLINE	Nature 409 (6821), 685-690 (2001)		
PUBMED	2105660		
REFERENCE	11217851		

REFERENCE
AUTHORS

5 (bases 1 to 1612)
Aachai,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A.,
Arakawa,T., Baldarelli,R., Bono,H., Brownstein,M., Bulc,C.,
Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T.,
Hara,A., Hayata,N., Hill,D., Hiramoto,K., Hiraoka,T., Horii,F.,
Hume,D., Imotani,K., Ishii,Y., Itoh,M., Iwama,M., Kasukawa,T.,
Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S.,
Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K.,
Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Ozakenbush,J.,
Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Saeki,D.,
Schirral,L., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,
Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yamanaka,I.,
Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M., and
Hayashizaki,Y.

JOURNAL

TITLE
Direct Submission
Submitted (10-JUL-2000) Yoshihide Hayashizaki. The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration and Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)

COMMENT
Please visit our web site (<http://genome.gsc.riken.go.jp/>) for
further details.
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues. First strand cDNA was primed with a primer
[5' GAGACGAGAAGCGCCGCACACTCGATCTTTTATTATTTTATVN 3'], cDNA was
prepared by using trihalose thermo-activated reverse transcriptase
and subsequently enriched for full-length by cap-trapper. cDNA went
through one round of normalization to Kot = 5.0. Second strand cDNA
was prepared with the primer adapter of sequence [5'
GAGACGAGAAGCATTCAGAGCTCAATTAATTAATAAACCCCCCCC 3']. cDNA was
cleaved with XhoI and SstI. Cloning sites, 5' end: SstI; 3' end:
XhoI. Host: SOLR.

FEATURES
source

Location/Qualifiers
1..1612
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="FANTOM DB:0610009G22"
/db_xref="MGD:MGI:1896762"
/db_xref="taxon:10090"
/clone="0610009G22"
/sex="male"
/tissue type="kidney"
/clone.lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
70..680
/note="data source:SPT, source key:O91U3, evidence:ISS
homolog to HYPOTHETICAL 15.3 KDA PROTEIN (FRAGMENT)
putative"
/codon_start=1
/protein_id="BAB2066.1"
/db_xref="GI:12832348"
/db_xref="MGD:MGI:1913292"
translation="MLPPAIHLSLPLCLIMRNCLAFKDATEILYSHVVPAPHP
SSNLTIANRGGRHFSTGLDRNSRVQVGRELRSKYISDGCTSIPIKELVCGA
ECLEPLFLPMWIGGGYGTKYSRSSDEMVCVDNKTKTORLOCOGDSRTYKIVTV
TACKCKTRYROHNSSINFESVSAPRAKHRRREKRASKSKLS"
1593..1598
/note="putative"
1612
/note="putative"

polyA_site
BASE COUNT 435 a 397 C 352 g 428 t
ORIGIN

Query Match 11.1%, Score 84.2; DB 11; Length 1612;
Best Local Similarity 52.7%; Pred. No. 1.4e-07;
Matches 288; Conservative 0; Mismatches 238; Indels 21; Gaps 4;

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 28, 2003, 17:20:23 ; Search time 237.766 Seconds

(without alignments)
7188.873 Million cell updates/sec

Title: US-09-867-274-1

Perfect score: 759
Sequence: 1 taccgaagcgcgcgtgcccc.....ctgcctctcgcgcgtggttc 759

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

N Geneseq.101002:*
1: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	759	100.0	759	24	ABA94293
2	759	100.0	2301	21	AAA29055
3	757.4	99.8	2301	21	AAA29061
4	757.4	99.8	2301	21	AAA29062
5	757	99.7	2329	22	AAA91023
6	757	99.7	2329	24	ABK69992
7	755.8	99.6	2301	21	AAA29056
8	735	96.8	2271	24	AA027576
9	642	84.6	642	21	AAA94051

10	613.2	80.8	642	21	AAA29057	Verter TGF-beta bi
11	536.2	70.6	674	21	AAA29059	Rat TGF-beta bindi
12	516.8	68.1	5680	21	AAA94049	Human DAN/Cerberus
13	516.8	68.1	7099	24	AA027577	Human osteolevin g
14	516.8	68.1	9301	21	AAA29064	Human TGF-beta bin
15	502	66.1	638	21	AAA29058	Murine TGF-beta bi
16	501.6	66.1	532	24	ABA94294	Mouse cloaked-2 po
17	467.8	61.6	532	21	AAA29060	Bovine TGF-beta bi
18	422	55.6	422	24	AAK45509	Human bone marrow
19	422	55.6	422	22	AB519777	Human genome-deriv
20	422	55.6	1104	21	AAA94050	Human DAN/Cerberus
21	372.2	49.0	35828	21	AAA29063	Murine TGF-beta bi
22	355	46.8	392	22	AAK32487	Human bone marrow
23	355	46.8	392	22	AB507272	Human genome-deriv
24	319.6	42.1	555	24	AB035254	Oligonucleotide fo
25	319.6	42.1	555	24	AB035255	Oligonucleotide fo
26	276.6	36.4	555	24	AB035256	Oligonucleotide fo
27	276.6	36.4	555	24	AB035257	Oligonucleotide fo
28	182.8	24.1	198	20	AAK40842	Human secreted pro
29	182.4	24.0	196	21	AAK05741	Human secreted pro
30	82.6	10.9	954	22	AAK9782	Skin cell cDNA, SE
31	82.6	10.9	954	24	AB134934	Murine cDNA isolat
32	82.6	10.9	962	21	AAZ61645	cDNA encoding mur
33	82.6	10.9	962	21	AAZ61736	cDNA encoding mur
34	82.6	10.9	962	22	AAK99578	Skin cell cDNA, SE
35	82.6	10.9	962	22	AAK99578	Skin cell cDNA, SE
36	82.6	10.9	962	24	AB134730	Murine cDNA isolat
37	82.6	10.9	962	24	AB134821	Murine cDNA isolat
38	69.6	9.2	618	24	ABK29025	Human breast tumou
39	69.6	9.2	621	24	ABK29024	Human breast tumou
40	69.6	9.2	900	18	AAK47661	cDNA encoding smal
41	69.6	9.2	900	19	AAV37735	Human small CCN-11
42	69.6	9.2	900	24	AB151660	Human small CCN-11
43	69.6	9.2	1091	24	AB159595	Human angiogenesis
44	69.6	9.2	1091	24	AB188106	Human PMS52 cDNA
45	69.6	9.2	1692	21	AAK79467	cDNA sequence of h

ALIGNMENTS

RESULT 1	ABA94293	standard; cDNA; 759 BP.
ID	ABA94293	
XX	ABA94293	
AC	26-MAR-2002	(first entry)
DT	Human cloaked-2	polypeptide encoding cDNA.
XX		
DE		
XX		
KW	Cloaked-2; cysteine knot motif; nephrotropic; cardiant; immunomodulator;	
KW	hepatotropic; antinflammatory; antithyroid; cytosolic; neuroprotective;	
KW	antianemic; hypotensive; antirhythmic; antidiabetic; anorectic; gene therapy; cell therapy; antisense therapy;	
KW	human; ss.	
XX		
OS	Homo sapiens.	
XX		
Key		Location/Qualifiers
FT	CDS	37..678
FT		/*tag= a
FT		/product= "cloaked-2 polypeptide"
FT	sig_peptide	37..105
FT		/*tag= b
FT	mat_peptide	106..759
FT		/*tag= c
FT		/note= "see ABB07207"
XX		
PN	WO200192308-A2.	
XX		
XX	06-DEC-2001.	
PD		
XX		

PF 29-MAY-2001; 2001MO-US17478.
 XX
 PR 01-JUN-2000; 2000US-208550P.
 PR 04-AUG-2000; 2000US-223542P.
 PA
 XX (AMGE-) AMGEN INC.
 PI Paszcy CJ, Gao Y;
 XX
 DR WPI: 2002-114325/15.
 DR P-PSDB; ABB07207, ABB07209.
 XX
 PT New human and mouse cystine-knot polypeptide designated as Cloaked-2,
 PT for treating or preventing kidney, heart (e.g. myocardial infarction)
 PT or liver (e.g. hepatitis) diseases -
 XX
 PS Claim 1; Fig 1; 170pp; English.
 XX
 CC The invention relates to polypeptides comprising a cysteine knot motif
 CC and designated as Cloaked-2, derived from human and mouse. The cloaked-2
 CC polypeptides can be expressed by standard recombinant methodology. The
 CC cloaked-2 polynucleotides are useful in gene therapy and antisense
 CC therapy. The cloaked-2 polypeptides and polynucleotides are useful for
 CC treating, preventing, ameliorating or detecting diseases and disorders of
 CC the kidney (e.g. anemia, hypertension or low blood pressure), heart (e.g.
 CC cardiac hypertrophy, congestive heart failure, myocardial infarction,
 CC arrhythmias, atherosclerosis, hypertension or low blood pressure),
 CC skeletal muscle (e.g. muscular dystrophy or cachexia), placenta (e.g.
 CC congenital abnormalities or miscarriage), liver (e.g. hepatitis or
 CC cirrhosis), pancreas (e.g. diabetes or pancreatitis), thyroid (e.g.
 CC Grave's disease or myxedema) or adrenal cortex (e.g. Cushing's disease
 CC or Addison's disease), homeostasis or metabolic diseases (e.g. obesity,
 CC cancer or myopathies), infections, or autoimmune diseases. Selective
 CC binding agents may be used to modulate the biological activities of
 CC cloaked-2 polypeptides or to detect Cloaked-2 polypeptide levels in a
 CC sample. Transgenic non-human animals are useful for drug candidate
 CC screening. The present sequence represents a cDNA encoding the human
 CC cloaked-2 polypeptide.
 CC
 SO Sequence 759 BP; 125 A; 282 C; 244 G; 108 T; 0 other;

Query Match 100.0%; Score 759; DB 24; Length 759;
 Best Local Similarity 100.0%; Pred. No. 4,3e-138;
 Matches 759; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TACTGGAAGGTTGGTCCCTCCTCTGCTGGTGAACATGACATCCACCTGGCCCTGTGT 60
 DB 1 TACTGGAAGGTTGGTCCCTCCTCTGCTGGTGAACATGACATCCACCTGGCCCTGTGT 60
 QY 61 CTGCTGCTGCTGTGTACACAGAGCTTCCGTAGTGGAGGCGCAGGGGTGAGAGCG 120
 DB 61 CTGCTGCTGCTGTGTACACAGAGCTTCCGTAGTGGAGGCGCAGGGGTGAGAGCG 120
 QY 121 TTCAAGAAATATCCACGGAATATATCCCGAGCTGGAGAGTACCCCGAGCTTCAACG 180
 DB 121 TTCAAGAAATATGCAAGAAATATATCCCGAGCTGGAGAGTACCCCGAGCTTCAACG 180
 QY 181 GAGCTGGAAGAACAAAGACCATGAACCGGGCGAGAAAGAGGGGGCTTCCCAACAC 240
 DB 181 GAGCTGGAAGAACAAAGACCATGAACCGGGCGAGAAAGAGGGGGCTTCCCAACAC 240
 QY 241 CCCTTTGAGACCAAGACGTGTCCGAGTACAGTGGCGGAGCTGTGACTTCAACCCGCTAC 300
 DB 241 CCCTTTGAGACCAAGACGTGTCCGAGTACAGTGGCGGAGCTGTGACTTCAACCCGCTAC 300
 QY 301 GTGACCGATGGGCGGCTGCGCAGAGCGGCTCAACGAGCTGTGTCTCCGCGCAG 360
 DB 301 GTGACCGATGGGCGGCTGCGCAGAGCGGCTCAACGAGCTGTGTCTCCGCGCAG 360
 QY 361 TGGGGCCCGGCGGCTGCTGCAAGCGCATGGCGCGGCAAGTGTGCGCACTTAAT 420
 DB 361 TGGGGCCCGGCGGCTGCTGCAAGCGCATGGCGCGGCAAGTGTGCGCACTTAAT 420

QY 421 GGGCCGACCTTCGGTGATCCCGACCGCTACCGCGGAGCGCGTGTGAGCTGTGT 480
 DB 421 GGGCCGACCTTCGGTGATCCCGACCGCTACCGCGGAGCGCGTGTGAGCTGTGT 480
 QY 481 CCGGTTGTTGAGGCGCGCGCGCGCGCAAGGTGCGCTGTGAGCTGTGCAATGCAAG 540
 DB 481 CCGGTTGTTGAGGCGCGCGCGCGCGCAAGGTGCGCTGTGAGCTGTGCAATGCAAG 540
 QY 541 CGGCTCACCGGCTTCCAAACAGTCCGAGCTCAAGACTTCCGGGACCGAGCCGCTCG 600
 DB 541 CGGCTCACCGGCTTCCAAACAGTCCGAGCTCAAGACTTCCGGGACCGAGCCGCTCG 600
 QY 601 CGCAGAAAGGCGCGGAGCCCGCGCGCGCGCGCGAGCGCCAAAGCCAAACGAGCGGAG 660
 DB 601 CGCAGAAAGGCGCGGAGCCCGCGCGCGCGCGCGAGCGCCAAAGCCAAACGAGCGGAG 660
 QY 661 CTGGAAGACGCTTACTAGAGCCCGCGCGCGCGCTCCCAACGCGGCGCGCGCGCT 720
 DB 661 CTGGAAGACGCTTACTAGAGCCCGCGCGCGCGCTCCCAACGCGGCGCGCGCGCT 720
 QY 721 GAACCGGCT 759
 DB 721 GAACCGGCT 759

RESULT 2
 AAA29055 standard; cDNA; 2301 BP.
 AC AAA29055;
 XX
 DT 12-SEP-2000 (first entry)
 DE Human TGF-beta binding protein (hBET) cDNA.
 XX
 KW osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
 KW hBET; chromosome 17q12-21; gene therapy; antisense therapy; fracture;
 KW bone mineralization; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 48..689
 FT /tag= a
 FT /product= TGF-beta_binding_protein
 FT
 FN WO200032773-A1.
 XX
 PD 08-JUN-2000.
 XX
 PF 24-NOV-1999; 99WO-US27990.
 XX
 PR 27-NOV-1998; 98US-0110283.
 XX
 PA (DARW-) DARWIN DISCOVERY LTD.
 XX
 PI Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepker BW;
 PI Van Ness J, Winkler DG;
 XX
 DR WPI: 2000-412321/35.
 DR P-PSDB; AAY96429.
 XX
 PT Nucleic acids (1) encoding a transforming growth factor beta binding
 PT protein, useful for identifying agents for treating osteopenia,
 PT osteoporosis and fractures
 XX
 PS Claim 1; Page 114-115; 162pp; English.
 CC This cDNA encodes a human transforming growth factor-beta (TGF-beta)
 CC binding protein designated hBET. The hBET gene has been localized
 CC to the chromosome 17q12-21. The cDNA and protein may be used for
 CC prevention, treatment and diagnosis of diseases associated with
 CC inappropriate hBET expression. For example, they may be used to treat

CC disordere associated with decreased TGF-beta B expression. The cDNA or
CC vectors may be administered to treat diseases by rectifying mutations or
CC deletions in a patient's genome that affect the activity of BEER by
CC expressing inactive proteins or to supplement the patients own production
CC of BEER polypeptides. The nucleic acids may be used for recombinant
CC production of BEER, gene therapy, antisense therapy, as probes for
CC diagnostic assays and for functional studies. BEER may be used to raise
CC antibodies and for identification of BEER modulators. BEER antagonists
CC may be used to increase bone mineral content for the treatment of
CC disorders such as osteopenia, osteoporosis, fractures and other
CC disorders associated with low mineral content.

50 Sequence 2301 BP; 568 A; 635 C; 614 G; 484 T; 0 other;

Query Match	100.0%	Score 759	DB 21	Length 2301
Best Local Similarity	100.0%	Pred. NC	4.4e-138	
Matches 759	Conservative 0	Mismatches 0	Indels 0	Gaps 0

Qy	1	TACTGGAAAGTGGCGTCGCTCTCTTGAGCTGGTACATGAGAGCTCCCATGGCCCTGGT	60
Db	12	TACTGGAAAGTGGCGTGGCTCTCTTGAGCTGGTACATGAGAGCTCCCATGGCCCTGGT	71
Qy	61	CTGTGTCGCTCTGAGTACACAGACCTTCGTTAGTGGAGCCAGAGGTGGAGGCG	120
Db	72	CTGTGTCGCTCTGAGTACACAGACCTTCGTTAGTGGAGGCGCAGAGGTGGAGGCG	133
Qy	121	TTTCAAAATGATGCCACGGAAATCATCCCGAGTCGGAGAGTACCCGAGCTTCACCG	186
Db	132	TTTCAAAATGATGCCACGGAAATCATCCCGAGTCGGAGAGTACCCGAGCTTCACCG	191
Qy	181	GAGCTGGAATCAACAGACCATGAAACGGAGCGAGAACGAGAGGAGGCTCCCAACAC	240
Db	192	GAGCTGGAATCAACAGACCATGAAACGGAGCGAGAACGAGAGGAGGCTCCCAACAC	255
Qy	241	CCCTTTGAGACCAAAAGACGTGTCCGAGTACAGCTGCGCGAGCTGCATCTACCCGTAAC	300
Db	252	CCCTTTGAGACCAAAAGACGTGTCCGAGTACAGCTGCGCGAGCTGCATCTACCCGTAAC	311
Qy	301	GTGACCGATGGGCGCTGCTGCGACAGGCGCCAAAGCTCGGTACACGAGTGGTGTGCTCGGCGAG	366
Db	312	GTGACCGATGGGCGCTGCTGCGACAGGCGCCAAAGCTCGGTACACGAGTGGTGTGCTCGGCGAG	371
Qy	361	TGGAGGCGCGGCGCGCTGCTGCCAACGCGCATGGGCGCGCGGAGAGTGTGGGACCTAGT	422
Db	372	TGGAGGCGCGGCGCGCTGCTGCCAACGCGCATGGGCGCGCGGAGAGTGTGGGACCTAGT	431
Qy	421	GGGCGCCGACTTCGCTGCTGATCCCGACCGCTACCGGCGGAGGCGCGTGGAGCTGTGTTGT	480
Db	432	GGGCGCCGACTTCGCTGCTGATCCCGACCGCTACCGGCGGAGGCGCGTGGAGCTGTGTTGT	491
Qy	481	CCCGGTGTGAGGCGCGCGCGCGCGGAGGAGTGAGCTGTGGGCTCTGTGACAGTGCAG	540
Db	492	CCCGGTGTGAGGCGCGCGCGCGCGGAGGAGTGAGCTGTGGGCTCTGTGACAGTGCAG	551
Qy	541	CGGCTCACCCCGTTTCAACAACAGTCCGAGACTCAAGAGCTTGGGAGCGAGGCGCTCGG	600
Db	552	CGGCTCACCCCGTTTCAACAACAGTCCGAGACTCAAGAGCTTGGGAGCGAGGCGCTCGG	611
Qy	601	CCGCAAGAGGCGCGAGAACCGCGGCGCCCGGCGCGGAGGCGCCAAAGCCAACAGGCGGAG	660
Db	612	CCGCAAGAGGCGCGAGAACCGCGGCGCCCGGCGCGGAGGCGCCAAAGCCAACAGGCGGAG	671
Qy	661	CTGGAGAACGCTTACTAGAGCGCGCGCGCGCTTCCCAACCGGCGGAGCGCCCGGCTT	720
Db	672	CTGGAGAACGCTTACTAGAGCGCGCGCGCGCTTCCCAACCGGCGGAGCGCCCGGCTT	731
Qy	721	GAAACCGGCGCCCACTTTCGTCTGTCTGTGGGCGGTGTGT 759	
Db	732	GAAACCGGCGCCCACTTTCGTCTGTCTGTGGGCGGTGTGT 770	

RESULT 3
AAA29061

ID	AAA29061standard; cDNA; 2301 BP.
XX	
AC	AAA29061;
XX	
XX	12-SEP-2000 (first entry)
XX	
DE	Mutant human TGF-beta binding protein (BBER) cDNA.
XX	
XX	osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
KW	BBER; chromosome 17q12-21; gene therapy; antisense therapy; fracture;
KW	bone mineralization; mutant; sclerosteosis; ss.
XX	
OS	Homo sapiens.
XX	
FT	Key
FT	location/Qualifiers
FT	CDS
FT	48..119
FT	/*tag= a
FT	/note= "mutant BBER created by sclerosteosis
FT	nonsense mutation"
XX	
PN	WO200032773-A1.
XX	
PD	08-UN-2000.
XX	
PF	24-NOV-1999; 99WO-US27990.
XX	
PR	27-NOV-1998; 98US-0110283.
XX	
PA	(DARW-) DARWIN DISCOVERY LTD.
PI	Brunkow ME, Gales DJ, Kovacevich B, Mulligan JT, Paepfer BW;
PI	Van Ness J, Winkler DG;
XX	
DR	WPI; 2000-412321/35.
XX	
DR	P-PSDB; AAY96435.
PT	Nucleic acids (1) encoding a transforming growth factor beta binding
PT	protein, useful for identifying agents for treating osteopenia,
PT	osteoporosis and fractures
XX	
PS	Claim 27; Page 117-118; 162pp; English.
XX	
CC	This cDNA encodes a mutant human transforming growth factor-beta
CC	(TGF-beta) binding protein (BBER) produced as a result of a nonsense
CC	mutation in the BBER coding sequence (C to T mutation at position 117).
CC	This mutation has been linked to sclerosteosis in humans by linkage
CC	analysis of an Afrikaner population in South Africa. The hBBER gene has
CC	been localized to the chromosome 17q12-21. The cDNA and protein may be
CC	used for prevention, treatment and diagnosis of diseases associated with
CC	inappropriate BBER expression. For example, they may be used to treat
CC	disorders associated with decreased TGF-beta BP expression. The cDNA or
CC	vectors may be administered to treat diseases by rectifying mutations or
CC	deletions in a patient's genome that affect the activity of BBER by
CC	expressing inactive proteins or to supplement the patients own production
CC	of BBER polypeptides. The nucleic acids may be used for recombinant
CC	production of BBER, gene therapy, antisense therapy, as probes for
CC	diagnostic assays and for functional studies. BBER may be used to raise
CC	antibodies and for identification of BBER modulators. BBER antagonists
CC	may be used to increase bone mineral content for the treatment of
CC	disorders such as osteopenia, osteoporosis, fractures and other
CC	disorders associated with low mineral content.
XX	
XX	
SQ	Sequence 2301 BP; 568 A; 634 C; 614 G; 485 T; 0 other;
	Query Match 99.8%; Score 757.4; DB 21; Length 2301;
	Best Local Similarity 99.9%; Pred. No. 8.9e-138;
	Matches 758; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY	1 TACTGGAAGTGGCGCTGCTCTCTGCGTGTATACATGACAGCTCCCACTGGCCCTGTGT 60
DB	12 TACTGGAAGTGGCGCTGCTCTCTGCGTGTATACATGACAGCTCCCACTGGCCCTGTGT 71
OY	61 CTCGTGCTGCTGCTGATACACAGACCTTCCTCGTGTAGTGAAGAGCCAGAGGTGACAGGC 120

Db	72	CTCGTCGCTGCTGCTGTCACACAGCCTTCCGTGTAGTGGAGGGCTAAAGGGGTGGACGGCG	131
Qy	121	TTCAAGAAATGATGCGACGGAATCATCTCCGAGCTCGGAGAGTAACTCCCGAGCTCTCAACCG	180
Db	132	TTCAAGATGATGTGCACAGGAATCATCTCCGAGCTCGGAGAGTAACTCCCGAGCTCTCAACCG	191
Qy	181	GAGCTGAGGAACAACAAGACCATGAAACCGGACGAGAAACGAGAGGGAGCTCCCAACAC	240
Db	192	GAGCTGAGGAACAACAAGACCATGAAACCGGACGAGAAACGAGAGGGAGCTCCCAACAC	251
Qy	241	CCCTTTGAGACCAAGAAGCTGTCCGAGTAACTGTCCCGAGCTGCACTTCAACCGCTAC	300
Db	252	CCCTTTGAGACCAAGAAGCTGTCCGAGTAACTGTCCCGAGCTGCACTTCAACCGCTAC	311
Qy	301	GTCAGCCATGGGCGGTGCGGACGAGGCCAAGCCGCTCAACGAGCTGGGTGCTCCGACAG	360
Db	312	GTCAGCCATGGGCGGTGCGGACGAGGCCAAGCCGCTCAACGAGCTGGGTGCTCCGACAG	371
Qy	361	TGCGGACCGGCGCGCTGTGCTGCCAAGCCCATCGGACCGGCAAGTGTGGCGAAGCTAGT	420
Db	372	TGCGGACCGGCGCGCTGTGCTGCCAAGCCCATCGGACCGGCAAGTGTGGCGAAGCTAGT	431
Qy	421	GGGCGCCGACTTTCGCTGATCCCCGACCGCTCAACGCGCGAGAGCGGTGAGCTGTGTGT	480
Db	432	GGGCGCCGACTTTCGCTGATCCCCGACCGCTCAACGCGCGAGAGCGGTGAGCTGTGTGT	491
Qy	481	CCCGTGTGTAGAGCGCGCGCGCGCAAGGTGTGCTGTGTGCTTGTGTCAAGTGCAG	540
Db	492	CCCGTGTGTAGAGCGCGCGCGCGCGCAAGGTGTGCTGTGTGCTTGTGTCAAGTGCAG	551
Qy	541	CGCTCACCGCTTTCACCAACGAGTCCGAGCTCAAGGACTTTCGAGACCGAGGCGCTCG	600
Db	552	CGCTCACCGCTTTCACCAACGAGTCCGAGGACTTTCGAGACCGAGGCGCTCG	611
Qy	601	CCGCAAGAGGGCGGAGACCGGCGCGCGCGCGCGAGACGCAAAAGCAACAGCGAG	660
Db	612	CCGCAAGAGGGCGGAGACCGGCGCGCGCGCGCGCGAGACGCAAAAGCAACAGCGAG	671
Qy	661	CTGGAAGACGCTTACTAGAGCGCGCGCGCGCTTCCCAACGCGCGGCGCGCGCGCT	720
Db	672	CTGGAAGACGCTTACTAGAGCGCGCGCGCGCGCTTCCCAACGCGCGGCGCGCGCGCT	731
Qy	721	GAAACCGGCGCGCAATTTCTGTCTCTCTGCGCGGT	759
Db	732	GAAACCGGCGCGCAATTTCTGTCTCTCTGCGCGGT	770
RESULT 4			
ID	AAA29062		
ID	AAA29062	standard; cDNA; 2301 BP.	
XX	AA	AAA29062;	
XX	AC		
XX	DT	12-SEP-2000 (first entry)	
XX	DE	Human TGF-beta binding protein (BEER) variant P38R cDNA.	
XX	DE	osteopathic; transforming growth factor-beta; TGF-beta; binding protein;	
KW	BEER; variant; P38R; gene therapy; antisense therapy; fracture;		
KW	Chromosome 17q12-21; bone mineralization; ss.		
XX	OS		
XX	OS	Homo sapiens.	
XX	Key	Location/Qualifiers	
FT	CDS	48..689	
FT	FT	/*tag= a	
FT	FT	/label= BEER variant P38R	
XX	PN	/product= TGF-beta_binding_protein	
XX	PD	WO200032773-A1.	
XX	PD	08-JUN-2000.	

24-NOV-1999; 99WO-US27990.
27-NOV-1998; 98US-0110283.
(DARW-) DARWIN DISCOVERY LTD.
Brunkow ME, Gals DJ, Kovacevich B, Mulligan JT, Paepfer BW;
Van Ness J, Winkler DG;
WPI; 2000-412321/35.
P-PsDB; AAY96436#.
Nucleic acids (1) encoding a transforming growth factor beta binding
protein, useful for identifying agents for treating osteopenia,
osteoporosis and fractures
Claim 27; Page 120-121, 162pp; English.

This cDNA encodes a variant human transforming growth factor- β (TGF- β) binding protein designated BEER P388. The encoded protein comprises a substitution of arginine for the wild-type proline at residue 38. The cDNA and protein may be used for prevention, treatment and diagnosis of diseases associated with inappropriate BEER expression. For example, they may be used to treat disorders associated with decreased TGF- β BP expression. The cDNA or vectors may be administered to treat diseases by rectifying mutations or deletions in a patient's genome that affect the activity of BEER by expressing inactive proteins or to supplement the patients own production of BEER polypeptides. The nucleic acids may be used for recombinant production of BEER, gene therapy, antisense therapy, as probes for diagnostic assays and for functional studies. BEER may be used to raise antibodies and for identification of BEER modulators. BEER antagonists may be used to increase bone mineral content for the treatment of disorders such as osteopenia, osteoporosis, fractures and other disorders associated with low mineral content.

Query Match	99.8%;	Score 757.4;	DB 21;	Length 230;
Best Local Similarity	99.9%;	Pred. No. 8.9e-138;		
Matches 758; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;

OY	1	TATGGAAAGTGGGGTGGCCCTCCCTGGAGTGTACAGTAGAGCTCCCACTGGCCCTTGT	60
Db	12	TACTGAAAGTGGGGTGGCCCTCCTCTGGCTGTATCCATGTCAAGCTTCCACTTGGCCCTTGT	71
OY	61	CTCGTCTGCTCTGTGTATCACACAGCCTTCCGTGTAGTGGAGGCCAGGGGTGGCAGCG	120
Db	72	CTCGTCTGCTCTGTGTATCACACAGCCTTCCGTGTAGTGGAGGCCAGGGGTGGCAGCG	131
OY	121	TTCAAGAAATGATGCCACGGAATCATCCCCGAGCTCGAGAGTACCCCGAGCTTCCACG	180
Db	132	TTCAAGAAATGATGCCACGGAATCATCCCGAGCTCGAGAGTACCCCGAGCTTCCACG	191
OY	181	GAGCTGGAACACACAGACCATGAACGGGCGGAGAACGAGGGGGGGCTTCCCAACAC	240
Db	192	GAGCTGGAACACACAGACCATGAACGGGCGGAGAACGAGGGGGGGCTTCCCAACAC	251
OY	241	CCCTTTGAGACCAAGAGGTGTCCGAGTACAGCTGGCGGAGCTGCACCTTCAACCGCTAC	300
Db	252	CCCTTTGAGACCAAGAGGTGTCCGAGTACAGCTGGCGGAGCTGCACCTTCAACCGCTAC	311
OY	301	GTGACCGATGGGGCGTGCAGCAGCGCCAAAGCCGGTCAACGAGCTGTGTCTCCGGCAG	360
Db	312	GTGACCGATGGGGCGTGCAGCAGCGCCAAAGCCGGTCAACGAGCTGTGTCTCCGGCAG	371
OY	361	TTCGGCCCGGCGCGCTGTCTGTGCCAAGCCCATGTGGCGGCGCAAGTGGTGGCACTAAT	420
Db	372	TTCGGCCCGGCGCGCTGTCTGTGCCAAGCCCATGTGGCGGCGCAAGTGGTGGCACTAAT	431
OY	421	GGGCGCGAATTCCGTGCATCCCGAAGCGTACCGGCGGCAAGCGGTGCACTGTCTGT	480


```
Db 312 GTGACCGATGGGCGGTGCGGAGCGCCAGCCGAGCTGAGTGTCTCCGCGCAG 371
Qy 361 TGGCGCCCCGCGCGCTGCTGCTCCCAAGCCATCGGCGCGCGCAAGTGTGGCACTTACT 420
Db 372 TGGCGCGCGCGCGCTGCTGCTCCCAAGCCATCGGCGCGCGCAAGTGTGGCACTTACT 431
Qy 421 GGGCGCGCACTTCCGCTGATCCCGGAGCCGCTAAGCGCGAGCGCTGAGCTGTGTGT 480
Db 432 GGGCGCGCACTTCCGCTGATCCCGGAGCCGCTAAGCGCGAGCGCTGAGCTGTGTGT 491
Qy 481 CCGCGTGTGAGCGCGCGCGCGCGCGCAAGGTGCGCTGTGTGTGTGTGTGTGTGTGT 540
Db 492 CCGCGTGTGAGCGCGCGCGCGCGCGCAAGGTGCGCTGTGTGTGTGTGTGTGTGTGT 551
Qy 541 CCGCGTGTGAGCGCGCGCGCGCGCGCAAGGTGCGCTGTGTGTGTGTGTGTGTGTGT 600
Db 552 CCGCGTGTGAGCGCGCGCGCGCGCGCAAGGTGCGCTGTGTGTGTGTGTGTGTGTGT 611
Qy 601 CCGCGAAGGCGCGGAGCGCGCGCGCGCGCGCAAGGTGCGCTGTGTGTGTGTGTGTGT 660
Db 612 CCGCGAAGGCGCGGAGCGCGCGCGCGCGCGCAAGGTGCGCTGTGTGTGTGTGTGTGT 671
Qy 661 CTGGAGAACGCTTACTAGAGCGCGCGCGCGCGCTTCCCAAGCGCGCGCGCGCGCTT 720
Db 672 CTGGAGAACGCTTACTAGAGCGCGCGCGCGCGCTTCCCAAGCGCGCGCGCGCGCTT 731
Qy 721 GAACCGCGCGCGCGCACTTCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 759
Db 732 GAACCGCGCGCGCGCACTTCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 770
```

RESULT 2

```
US-09-449-218D-3
; Sequence 3, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepker, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: BONE MINERALIZATION
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 2301
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-449-218D-3
```

```
Query Match 99.8%; Score 757.4; DB 4; Length 2301;
Best Local Similarity 99.9%; Pred. No. 1e-144;
Matches 758; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 TACTGGAAGTGGCGTGCCTCTCTGTGCTGTAACATGAGCTCCCACTGAGCTGTGT 60
Db 12 TACTGGAAGTGGCGTGCCTCTCTGTGCTGTAACATGAGCTCCCACTGAGCTGTGT 71
Qy 61 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
Db 72 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 131
Qy 121 TTCAAGATGATGCGACGGAATCATCCCGAGCTCGAGAGTACCCCGAGCTTCCACCG 180
Db 132 TTCAAGATGATGCGACGGAATCATCCCGAGCTCGAGAGTACCCCGAGCTTCCACCG 191
```

```
Qy 181 GAGCTGGAAGCAACAGACCATGATACCGGCGGAGAGAACGAGGCGCGCTCCCAACAC 240
Db 192 GAGCTGGAAGCAACAGACCATGATACCGGCGGAGAGAACGAGGCGCGCTCCCAACAC 251
Qy 241 CCGTTTGAAGCAACAGAGCTGTGAGTACAGCTGCGCGAGCTGCACTTACCCGCTAC 300
Db 252 CCGTTTGAAGCAACAGAGCTGTGAGTACAGCTGCGCGAGCTGCACTTACCCGCTAC 311
Qy 301 GTGACGATGGGCGGTGCGCGAGCGCGCAAGCGGCTCAAGCTGATGTGTGTGTGTGTGT 360
Db 312 GTGACGATGGGCGGTGCGCGAGCGCGCAAGCGGCTCAAGCTGATGTGTGTGTGTGTGT 371
Qy 361 TGGCGCGCGCGCGCTGCTGCTCCCAAGCCATCGGCGCGCAAGTGTGTGTGTGTGTGT 420
Db 372 TGGCGCGCGCGCGCTGCTGCTCCCAAGCCATCGGCGCGCAAGTGTGTGTGTGTGTGT 431
Qy 421 GGGCGCGCACTTCCGCTGATCCCGGAGCCGCTAAGCGCGAGCGCGCTGTGTGTGTGT 480
Db 432 GGGCGCGCACTTCCGCTGATCCCGGAGCCGCTAAGCGCGAGCGCGCTGTGTGTGTGT 491
Qy 481 CCGGCTGTGAGCGCGCGCGCGCGCGCAAGGTGCGCTGTGTGTGTGTGTGTGTGTGT 540
Db 492 CCGGCTGTGAGCGCGCGCGCGCGCGCAAGGTGCGCTGTGTGTGTGTGTGTGTGTGT 551
Qy 541 CCGCTCACCGCGCTTCCACACACAGTGGAGCTCAAGACTTGGGAGCCGAGCGCTTCCG 600
Db 552 CCGCTCACCGCGCTTCCACACACAGTGGAGCTCAAGACTTGGGAGCCGAGCGCTTCCG 611
Qy 601 CCGCGAAGGCGCGGAGCGCGCGCGCGCGCGCGCGCGCGCAAGGTGCGCTGTGTGTGT 660
Db 612 CCGCGAAGGCGCGGAGCGCGCGCGCGCGCGCGCGCGCGCAAGGTGCGCTGTGTGTGT 671
Qy 661 CTGGAGAACGCTTACTAGAGCGCGCGCGCGCGCTTCCCAAGCGCGCGCGCGCGCTT 720
Db 672 CTGGAGAACGCTTACTAGAGCGCGCGCGCGCGCTTCCCAAGCGCGCGCGCGCGCTT 731
Qy 721 GAACCGCGCGCGCGCACTTCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 759
Db 732 GAACCGCGCGCGCGCACTTCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 770
```

RESULT 3

```
US-09-449-218D-7
; Sequence 7, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepker, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: BONE MINERALIZATION
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 2301
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-449-218D-7
```

```
Query Match 99.8%; Score 757.4; DB 4; Length 2301;
Best Local Similarity 99.9%; Pred. No. 1e-144;
Matches 758; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 TACTGGAAGTGGCGTGCCTCTCTGTGCTGTAACATGAGCTCCCACTGAGCTGTGT 60
Db 12 TACTGGAAGTGGCGTGCCTCTCTGTGCTGTAACATGAGCTCCCACTGAGCTGTGT 71
```


RESULT 8
US-09-449-218D-18
; Sequence 18, Application US/09449218D
; Patent No. 6395511

QY	208	CGGGCGGAGAACGAGAGGCGCGCTCCCAACCAACCTTTTGAGACCAAAACCTGTCCGAG	26
Dp	3138	CAGGGGTGGCCAGCGCGGGCCACCCTCAACGCGGCTCTCTCCACAGACGTGTCCAG	319
QY	268	TACAGCTGCGCGAGAGTGCATCTTCAACCCGCTACGTAGCAACGATGGGCGTGGCCAGCGCC	327
Dp	3198	TACAGCTGCGCGAGAGTGCATCTTCAACCCGCTACGTAGCAACGATGGGCGTGGCCAGCGCC	325
QY	328	AAGCGCGGTACCGAGAGTGGAGTGCCTCGGCGAGTGGGCGCGCGCGCTGTGTCGCAAC	387
Dp	3258	AAGCGCGGTACCGAGAGTGGAGTGCCTCGGCGAGTGGGCGCGCGCGCTGTGTCGCAAC	331
QY	388	GCCATCGCGCGCGGCAAGTGGTGGCCAGCTTAGTGGGCGCGGATTCGCGCTGATGCCGAC	447
Dp	3318	GCCATCGCGCGCGGCAAGTGGTGGCCAGCTTAGTGGGCGCGGATTCGCGCTGATGCCGAC	337
QY	448	CGCTAACCGCGCGAGCGCGTGCAGCTGTGTGCCGAGTGGTAGAGCGCGCGCGCGCGC	507
Dp	3378	CGCTAACCGCGCGAGCGCGTGCAGCTGTGTGCCGAGTGGTAGAGCGCGCGCGCGCGCGC	343
QY	508	AAGGTGGCGCTGGTGGCTCTGTGTCAAGTGCAGAGCGCTTACCCGCTTCCACAACCAAGTCG	567
Dp	3438	AAGGTGGCGCTGGTGGCTCTGTGTCAAGTGCAGAGCGCTTACCCGCTTCCACAACCAAGTCG	349
QY	568	GAGCTCAAGGACTTTCGGGAGCCGAGGCGCGCTCGGCGCGCAAAAGGCGGAAAGCGCGGCGCC	627
Dp	3498	GAGCTCAAGGACTTTCGGGAGCCGAGGCGCGCTCGGCGCGCAAAAGGCGGAAAGCGCGGCGCC	355
QY	628	CGCGCCCGAGCGCCCAAAGCCAAACAGGCGAGGCTGTGAGAAAGCTTACGTAGAGCCCGGCC	687
Dp	3558	CGCGCCCGAGCGCCCAAAGCCAAACAGGCGAGGCTGTGAGAAAGCTTACGTAGAGCCCGGCC	361
QY	688	GGCGCCCTCCCCACCGCGCGGAGCGCCCGGCGCTGAAACCGCGCGCCCAATTTCTGTCTTC	747
Dp	3618	GGCGCCCTCCCCACCGCGCGGAGCGCCCGGCGCTGAAACCGCGCGCCCAATTTCTGTCTTC	367
QY	748	TGGCGGTGGTTTT	759
Dp	3678	TGGCGGTGGTTTT	3689

RESULT 9
US-09-449-218D-11
; Sequence 11, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.

```

1  APPLICANT: Kovacevich, Brian
2  APPLICANT: Mulligan, John T.
3  APPLICANT: Paepfer, Bryan W.
4  APPLICANT: Van Ness, Jeffrey
5  APPLICANT: Winkler, David G.
6  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
7  TITLE OF INVENTION: BONE MINERALIZATION
8  FILE REFERENCE: 240083.508
9  CURRENT APPLICATION NUMBER: US/09/449,218D
10 CURRENT FILING DATE: 1999-11-24
11 NUMBER OF SEQ ID NOS: 45
12 SOFTWARE: FastSeq for Windows Version 3.0
13 SEQ ID NO 11
14 LENGTH: 638
15 TYPE: DNA
16 ORGANISM: Mus musculus
17 US-09-449-218D-11

```

```
Query Match 66.1%; Score 502; DB 4; Length 638;
Best Local Similarity Pred No 3e-93;
Matches 563; Conservative 0; Mismatches 75; Indels 6; Gaps 1.
```

[illegible]

RESULT 10
US-09-449-218D-15
; Sequence 15, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.

```

1  APPLICANT: Galas, David U.
2  APPLICANT: Kovacevich, Brian
3  APPLICANT: Mulligan, John T.
4  APPLICANT: Paepker, Bryan W.
5  APPLICANT: Van Nese, Jeffrey
6  APPLICANT: Winkler, David G.
7  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
8  TITLE OF INVENTION: BONE MINERALIZATION
9  FILE REFERENCE: 240083.508
10 CURRENT APPLICATION NUMBER: US/09/449,218D
11 CURRENT FILING DATE: 1999-11-24
12 NUMBER OF SEQ ID NOS: 45
13 SOFTWARE: FastSeq for Windows Version 3.0
14 SEQ ID NO 15
15 LENGTH: 532
16 TYPE: DNA
17 ORGANISM: Bos taurus
18 US-09-449-218D-15

```

Query Match:	61.6%;	Score 467.8;	DB 4;	Length 532;
Best Local Similarity	-93.5%;	Pred. No. 2.4e-86;		
Matches 500;	Conservative 0;	Mismatches 32;	Indels 3;	Gaps 1;

QY	125	AGAAAGATGCAAGAAATCATCCCGAGCTCCGAGAGTACCCCGAGCTTCCACCCGAGC	184
Db	1	AGATGATGCCACAGAAATCATCTCCCGAGCTGGCGAGTACCCCGAGCTCTTCGCAAGAC	60
QY	185	TGAGAAACAAACAAGCATTAACCGGAGCGAGAACGAGAGGCGGACTTCCCAACCACTCCCT	244
Db	61	T---GAACAAACAAGACATTAACCGGAGCGAGAACGAGAGGAGACCTTCCCAACCACTCCCT	117
QY	245	TTTGAGACCAAGACGATGTCGAGTACAGTGGCCGAGAGTCACTTACCCGCTACGTA	304
Db	118	TTTGAGACCAAGACGCTCCGAGTACAGTGTCCGAGAGCTGCATTACCCGCTACGTA	177
QY	305	CCGATGGGCGGTGCGCGACGCGCAAGCGGATCACCGAGCTGTGTCTCCGACCACTGGG	364
Db	178	CCGATGGGCGGTGCGCGACGCGCAAGCGGATCACCGAGCTGTGTCTCCGAGCAGTGGG	237
QY	365	GCCCGGCGGCGCTGCTGCGCCCAACGCGCATGCGCGCGCAATGTGTGCGACCTTAGTGGC	424
Db	238	GCCCGGCGGCGCTGCTGCGCCCAACGCGCATGCGCGCGCAATGTGTGCGCCCAACGCGGC	297
QY	425	CCGACCTTCGCTGCATCCCGACCGGCTACCGGCGGACGCGTGCAGCTCTGTGTCCG	484
Db	298	CCGACTTCCGCTGCATCCCGACCGGCTACCGGCGGCGGCTGCAGCTCTGTGTCTGTG	357
QY	485	GTTGTGAGGCGCGCGCGGCGGCGAAGGTGCGCTGTGTGGCTCGTGCAGTGCAGGCGCC	544
Db	358	GCGGCGGCGGCGCGCGGCGGCGGCGAAGGTGCGCTGTGTGGCTCGTGCAGTGCAGGCGCC	417
QY	545	TCAACCGCTTCCACAACCAAGTGGAGCTTCAAGGACTTTCGGGACCGAGGCGGCTTTCGCGCG	604
Db	418	TCACTCGCTTCCACAACCAAGTGGAGCTTCAAGGACTTTCGGGACCGAGGCGGCGGCGCGC	477
QY	605	AGAGAGGCGCGGAAGCGCGGCGCCCGCGCCCGGACGCGCAAGGCCAACCGAGCGGA	659
Db	478	AAAGAGGCGCGGAAGCTGCGCGCCCGCGCCCGGAGCAACCAAGCCGCGGCGGA	532

RESULT 11
 US-09-449-218D-17
 ; Sequence 17, Application US/09449218D
 ; Patent No. 6395511
 ; GENERAL INFORMATION:
 ; APPLICANT: Brunkow, Mary E.
 ; APPLICANT: Galas, David J.
 ; APPLICANT: Kovacevich, Brian
 ; APPLICANT: Mulligan, John T.
 ; APPLICANT: Paepfer, Bryan W.
 ; APPLICANT: Van Ness, Jeffrey
 ; APPLICANT: Winkler, David G.
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING

```

; TITLE OF INVENTION: BONE MINERALIZATION
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 17
; LENGTH: 35828
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(35828)
; OTHER INFORMATION: n = A,T,C or G
US-09-449-218D-17

```

```

Query Match          49.0%; Score 372.2; DB 4; Length 35828;
Best Local Similarity 85.9%; Pred. No. 5.3e-67;
Matches 413; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

```

```

QY 222 AGGCGGCGCTCCCGCACCCTTTGAGACCAAGAGTGTCCAGATACAGTCCGCGA 281
DB 20988 AGGCTGACAGCCCTTCACGACATCTCTCCGAGATGTCCGAGTACAGTCCGCGA 21047
QY 282 GCTGCACTTACCCGCTACGTAACGATGGGCGGTGCGCGACCAAGCCGCTACCGA 341
DB 21048 GCTGCACTACACCCGCTTCTGTAAGACGCGCCATGCGCGACCAAGCCGCTACCGA 21107
QY 342 GCTGTGTGTCTCCGCGCAATGTCGCGCCGCGCGCTGTCTGCTCCCAAGCCATGCGCGG 401
DB 21108 GTTGTGTGTCTCCGCGCAATGTCGCGCCGCGCGCTGTCTGCTCCCAAGCCATGCGCGG 21167
QY 402 CAAGTGTGTGCACTAGTGTGGGCGCGCACTTCCGCTGATCCCGACCGCTACCGCGCA 461
DB 21168 GAAGTGTGTGCGCGCGCAAGCGCAATTTCCGCTGATCCCGACCGCTACCGCGCA 21227
QY 462 GCGGTGTGCACTGTGTGTCCGATGTGTGAGGCGCGCGCGCAAGTGTGCGCTGTGT 521
DB 21228 GCGGTGTGCACTGTGTGTCCGATGTGTGAGGCGCGCGCGCAAGTGTGCGCTGTGT 21287
QY 522 GCGGTGTGTGCAATGTGCAAGGCGCTTCAACCCGCTTCCCAACCAATGCGAGTCAAGACTT 581
DB 21288 GCGGTGTGTGCAATGTGCAAGGCGCTTCAACCCGCTTCCCAACCAATGCGAGTCAAGACTT 21347
QY 582 CGGACCGAGGCGCTGTGCGCGCGCAAGGCGCGAGGCGCGCGCGCGCGCGCGCGCG 641
DB 21348 CGGACCGAGGCGCTGTGCGCGCGCGCAAGGCGCGAGGCGCGCGCGCGCGCGCGCG 21407
QY 642 CAAGCCCAACGAGCGCGAGCTGAGAAAGCTTACAGAGCCCGCGCGCGCGCTTCCCAAC 701
DB 21408 CAAGCCCAACGAGCGCGAGCTGAGAAAGCTTACAGAGCCCGCGCGCGCGCTTCCCAAC 21467
QY 702 C 702
DB 21468 C 21468

```

```

RESULT 12
US-09-188-930-40
; Sequence 40, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James Greg
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; FILE REFERENCE: 11000.1011c1
; CURRENT APPLICATION NUMBER: US/09/188,930A
; NUMBER OF SEQ ID NOS: 348

```

```

; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 40
; LENGTH: 962
; TYPE: DNA
; ORGANISM: mouse
US-09-188-930-40

```

```

Query Match          10.9%; Score 82.6; DB 3; Length 962;
Best Local Similarity 52.5%; Pred. No. 1.1e-08;
Matches 287; Conservative 0; Mismatches 239; Indels 21; Gaps 4;

```

```

QY 112 TGGCAGCGCTTCAAGATGATGCGACGGAATCATCTCCGAGCTCGAGACTCCCGAG 171
DB 107 TGTGTGCTTTTAAAAAATGATGCGCACAGAAATCTTTTATTCATGTTGTTAAACTCTTC 166
QY 172 CCTCAGCGGAGCTGGAACAACAACGATGACCGGCGGAGAGAGGAGGCGGCT 231
DB 167 CCGGAC--ACCCAGCAGCAAGCAGCCTGATTAAGCCAGAGAGGAGGAGGAT 223
QY 232 CCCACCAACC--CTTGAAGCAAAAGCGTGTCCAGTACAGCTGCGCGAGCTGCAC 288
DB 224 TTCAATGACACTGAGCTGATGTAACAGTGAAGTTCAAGTGGCTGACGGAGACTGGG 283
QY 289 TTACCCCGCTACGTGACCGATGAGCGCTGTCCGACGCGCAAGCCGCTGACCGAGTGTG 348
DB 284 TCCACCAATATCATTTCCGAGCGCGCAGTGACACAGCATCAGCCCTGTAAGAGAGCTGTG 343
QY 349 TGCTCCGCGCAGTGTGCGCGCGCGCGCGCTGTGCTGCCAAGCCATGCGCGCGG-- 402
DB 344 TGCGGCGGAGTGTGCTTCCCTGCGGTGCTTCCCAATGATGAGAGAGCTACCGA 403
QY 403 -----AAGTGTGCGCACTAGTGTGGCGCGCACTTCCGCTGATCCCGACCGCTACCGC 456
DB 404 ACAAGTACTGTGAGCGCGAGAGAGCTCTCAGAGTGGCGGTGTGTAAGAAAGCGCGC 463
QY 457 GCGCAGCGGTGACGTCTGTGTCCGATGTGTGAGGCGCGCGCGCGCGAGTGTGCGC 516
DB 464 ACCCAGAGATCTCAGCTCAGTGTGAGAGCG--CAGCAGCGCGCACTTCAAAATCAAC 520
QY 517 CTGTGTGCGCTGTGCAATGTGCAAGGCGCTCAACCGCTTCCCAACCAATGCGAGTGTGAG 576
DB 521 GTGTGACGCGCTGTGCAATGTGCAAGGCTACCGCTTCAACCAATGCGAGTGTGCAAC 580
QY 577 GACTTGTGAGACCGAGCGCTGTGCGCGCGCAAGAGGCGCGAGCGCGCGCGCGCG 636
DB 581 TTTGAAGGTGTGTGCGCAGCGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 640
QY 637 AGCGCCA 643
DB 641 AATCCA 647

```

```

RESULT 13
US-09-188-930-209
; Sequence 209, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James Greg
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; FILE REFERENCE: 11000.1011c1
; CURRENT APPLICATION NUMBER: US/09/188,930A
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 209
; LENGTH: 962
; TYPE: DNA
; ORGANISM: Mouse

```


ORGANISM: Homo sapien
US-09-602-877A-92

Query Match 9.2%; Score 69.6; DB 4; Length 1692;
Best Local Similarity 50.0%; Pred. No. 4.7e-06;
Matches 239; Conservative 0; Mismatches 224; Indels 15; Gaps 2;

```
QY 110 GGTGGCAGGCGCTTCAAGATGATGCCAGGAATCATCCCGAGCTCGAGAGTACCCCG 169
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1582 GCTGTTTGGCTTTTAAAAATGATGCCAGGAATCTTTATTTCATGATGTTTAACTTG 1523
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 170 AGCCTCACCAGGAGCTGGAAGAACAAAGACATGAACCGGGCGAGAACGAGGGCGCG 229
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1522 TTCAGACACACCCAGACAGCAAGCAAGTTGAATCAAGCCAGAAATGAGGCGGCATT 1463
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 230 CTCCCCACCAACCCCTTTGAGACCAAGACGTGTCGAGTACAGCTGCGGAGCTGCACT 289
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1462 TCAGTACACTGAGACTGGATCGGAACACTCGGGTTCAAGTGGTGGCCGGAACTGCGTT 1403
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 290 TCACCCGCTACGTGACCGATGGGCCGTGCCGACGCGCAAGCCGGTCAACGAGCTGGTGT 349
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1402 CCACCAATATCATCTGTGATGGCCAGTGCACAGCATCAGCCCTGTGAGAGAGCTGGTGT 1343
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 350 GCTCCCGGCAAGTGGGCCCGCGCGCTGCTGCCCAAGCCCATGGGCCCGCGC----- 402
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1342 GTGCTGGCGAGGTGCTTGTCCCTGCAAGTCTCCTTAAGTGAAGAGGCTATAGAA 1283
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 403 -----AAGTGTGGCGACCTAGTGGGCCCGACTTCCGCTGCATCCCGACCGCTACCGCG 457
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1282 CAAAGTACTGAGAGAGAGAGAGTCCCAAGAGAGTGGCGGTGTCAATGACMAAAACCGTA 1223
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 458 CGCAGCGCGTGCAGCTGCTGTGTCGCGGTGAGGGCGCGCGCGCGCAAGGTGGCGC 517
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1222 CCCAGGAATTCAGCTCAGTGCAGGCAAGATG3---CAGCACACGCACTTCAAAATCACAG 1166
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 518 TGGTGGCTGTGTGCAAGTGAAGGCGCTCAACCGCTTCCAAACCAAGTTCGAGCTCAA 575
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1165 TAGTCACTGCTGCAAGTGAAGAGGTACACCGGACGACCAAGAGTTCAGTCAACA 1108
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: March 29, 2003, 01:30:43
Job time : 71.8796 secs

•
•
•
•

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 28, 2003, 14:14:27 ; Search time 11.6533 Seconds
(without alignments)
479.722 Million cell updates/sec

Title: US-09-867-274-2

Perfect score: 1049

Sequence: 1 QGMQAFKNDATETIIPELGEY.....KPRPARSAKNOALENNAY 190

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Issued Patents AA:*

2: /cgn2_6/prodata/1/1aa/5A_COMB.pep:*

3: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*

4: /cgn2_6/prodata/1/1aa/5A_COMB.pep:*

5: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*

6: /cgn2_6/prodata/1/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	* Match Length	DB	ID	Description
1	1049	100.0	213	4	US-09-449-218D-2
2	1040	99.1	213	4	US-09-449-218D-6
3	1040	99.1	213	4	US-09-449-218D-8
4	1033	98.5	213	4	US-09-449-218D-10
5	974	92.9	213	4	US-09-449-218D-14
6	951	90.7	211	4	US-09-449-218D-12
7	901.5	85.9	211	4	US-09-449-218D-16
8	344	32.8	206	4	US-09-188-930-159
9	344	32.8	206	4	US-09-188-930-286
10	341	32.5	206	1	US-08-468-847B-2
11	341	32.5	206	1	US-08-468-847B-20
12	89.5	8.5	184	4	US-09-449-218D-42
13	89.5	8.5	184	4	US-09-040-229B-2
14	87.5	8.3	184	4	US-09-040-229B-6
15	87.5	8.3	184	4	US-09-191-647-9
16	87.5	8.3	184	4	US-09-540-245A-9
17	87.5	8.3	184	4	US-09-540-153-9
18	86.5	8.2	182	4	US-09-040-229B-4
19	86.5	8.2	182	4	US-09-336-093-5
20	86	8.2	184	4	US-09-040-229B-9
21	85	8.1	180	4	US-09-449-218B-44
22	85	8.1	181	4	US-09-040-229B-8
23	81	7.7	267	4	US-08-878-474-1
24	80.5	7.7	267	4	US-09-449-218D-43
25	80	7.6	1480	3	US-09-191-647-7
26	80	7.6	1480	3	US-09-540-245A-7
27	80	7.6	1480	4	US-09-540-153-7

28	80	7.6	1480	4	US-09-182-024A-5	Sequence 5, Appl1
29	80	7.6	1480	5	PCT-US91-09055-2	Sequence 2, Appl1
30	79.5	7.6	830	4	US-09-562-737-39	Sequence 19, Appl1
31	79	7.5	336	4	US-08-804-166-8	Sequence 8, Appl1
32	79	7.5	336	4	US-08-910-991-8	Sequence 8, Appl1
33	78	7.4	1210	2	US-08-484-438-7	Sequence 7, Appl1
34	78	7.4	1210	2	US-08-475-035-4	Sequence 4, Appl1
35	77.5	7.4	510	5	PCT-US91-09055-9	Sequence 9, Appl1
36	77	7.3	510	4	US-08-246-489-2	Sequence 2, Appl1
37	74	7.1	470	2	US-08-946-241B-2	Sequence 2, Appl1
38	74	7.1	470	3	US-09-309-053-2	Sequence 2, Appl1
39	74	7.1	479	2	US-08-946-241B-9	Sequence 9, Appl1
40	74	7.1	479	3	US-09-309-053-9	Sequence 9, Appl1
41	73.5	7.0	685	2	US-08-878-989-1	Sequence 2, Appl1
42	73.5	7.0	685	3	US-09-136-282-2	Sequence 1, Appl1
43	73.5	7.0	685	4	US-09-272-796-1	Sequence 1, Appl1
44	73.5	7.0	685	4	US-09-505-744-2	Sequence 2, Appl1
45	73	7.0	301	3	US-08-303-861-21	Sequence 21, Appl1

ALIGNMENTS

```

RESULT 1
US-09-449-218D-2
; Sequence 2, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepfer, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449, 218D
; CURRENT FILING DATE: 1999-11-24
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 213
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-449-218D-2

Query Match          100.0%; Score 1049; DB 4; Length 213;
Best Local Similarity 100.0%; Pred. No. 1.4e-103;
Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGMQAFKNDATETIIPELGHPPEPELENKTMNFAENGRRPHHPFTKQVSYSCREL 60
DB 24 QGMQAFKNDATETIIPELGHPPEPELENKTMNFAENGRRPHHPFTKQVSYSCREL 83

QY 61 HETRYVTGPGCSAPRYVELVCSGCGPARLLPNAIGKMMRPSPGPRCTCPZYRAOR 120
DB 84 HETRYVTGPGCSAPRYVELVCSGCGPARLLPNAIGKMMRPSPGPRCTCPZYRAOR 143

QY 121 VOLLCPGSGAPPARKRVRLVASCCKRLTRFNQSELKDFTEAARPOKGRKPRPARSAK 180
DB 144 VOLLCPGSGAPPARKRVRLVASCCKRLTRFNQSELKDFTEAARPOKGRKPRPARSAK 203

QY 181 ANQALENNAY 190
DB 204 ANQALENNAY 213

RESULT 2
US-09-449-218D-6
; Sequence 6, Application US/09449218D
; Patent No. 6395511

```

```
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepfer, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 213
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-449-218D-6
```

```
Query Match          99.1%; Score 1040; DB 4; Length 213;
Best Local Similarity 99.5%; Pred. No. 1.3e-102;
Matches 189; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 OGWQAFKNDATETIIPELGEYPEPPPELENNKTNMRAENGGRPHHPETKDVSEYSCREL 60
DB 24 OGWQAFKNDATETIIPELGEYPEPPPELENNKTNMRAENGGRPHHPETKDVSEYSCREL 83
QY 61 HFTRYVTDGPCRSAPVTELVCSCGCGPARLLPNAIGRGKWMRPSGDPFCIPDRYRAOR 120
DB 84 HFTRYVTDGPCRSAPVTELVCSCGCGPARLLPNAIGRGKWMRPSGDPFCIPDRYRAOR 143
QY 121 VOLLCPGGEAPRARKVRLVASCKCKRLTRFNOSLKDFTGEARPOKGRPRPARSAK 180
DB 144 VOLLCPGGEAPRARKVRLVASCKCKRLTRFNOSLKDFTGEARPOKGRPRPARSAK 203
QY 181 ANQAELENAY 190
DB 204 ANQAELENAY 213
```

```
RESULT 3
US-09-449-218D-8
; Sequence 8, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepfer, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 213
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-449-218D-8
```

```
Query Match          99.1%; Score 1040; DB 4; Length 213;
Best Local Similarity 99.5%; Pred. No. 1.3e-102;
Matches 189; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
DB 24 OGWQAFKNDATETIIPELGEYPEPPPELENNKTNMRAENGGRPHHPETKDVSEYSCREL 83
QY 61 HFTRYVTDGPCRSAPVTELVCSCGCGPARLLPNAIGRGKWMRPSGDPFCIPDRYRAOR 120
DB 84 HFTRYVTDGPCRSAPVTELVCSCGCGPARLLPNAIGRGKWMRPSGDPFCIPDRYRAOR 143
QY 121 VOLLCPGGEAPRARKVRLVASCKCKRLTRFNOSLKDFTGEARPOKGRPRPARSAK 180
DB 144 VOLLCPGGEAPRARKVRLVASCKCKRLTRFNOSLKDFTGEARPOKGRPRPARSAK 203
QY 181 ANQAELENAY 190
DB 204 ANQAELENAY 213
```

```
RESULT 4
US-09-449-218D-10
; Sequence 10, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepfer, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 213
; TYPE: PRT
; ORGANISM: Cercopithecus pygerythrus
US-09-449-218D-10
```

```
Query Match          98.5%; Score 1033; DB 4; Length 213;
Best Local Similarity 98.4%; Pred. No. 6.9e-102;
Matches 187; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 1 OGWQAFKNDATETIIPELGEYPEPPPELENNKTNMRAENGGRPHHPETKDVSEYSCREL 60
DB 24 OGWQAFKNDATETIIPELGEYPEPPPELENNKTNMRAENGGRPHHPETKDVSEYSCREL 83
QY 61 HFTRYVTDGPCRSAPVTELVCSCGCGPARLLPNAIGRGKWMRPSGDPFCIPDRYRAOR 120
DB 84 HFTRYVTDGPCRSAPVTELVCSCGCGPARLLPNAIGRGKWMRPSGDPFCIPDRYRAOR 143
QY 121 VOLLCPGGEAPRARKVRLVASCKCKRLTRFNOSLKDFTGEARPOKGRPRPARSAK 180
DB 144 VOLLCPGGEAPRARKVRLVASCKCKRLTRFNOSLKDFTGEARPOKGRPRPARSAK 203
QY 181 ANQAELENAY 190
DB 204 ANQAELENAY 213
```

```
RESULT 5
US-09-449-218D-14
; Sequence 14, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepfer, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
US-09-449-218D-14
```


TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
FILE REFERENCE: 240083.508
CURRENT APPLICATION NUMBER: US/09/449,218D
CURRENT FILING DATE: 1999-11-24
NUMBER OF SEQ ID NOS: 45
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 14
LENGTH: 213
TYPE: PRT
ORGANISM: Rattus norvegicus
US-09-449-218D-14

Query Match 92.9%; Score 974; DB 4; Length 213;
Best Local Similarity 92.1%; Pred. No. 1.2e-95;
Matches 175; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 QGWAQKNDATETIIPELGEYEPPEPELENNKTMRANENGRPPHPEPEKDVSEYSCREL 60
DB 24 QGWAQKNDATETIIPGLREYEPPEPELENNKTMRANENGRPPHPPYDQVSEYSCREL 83
QY 61 HFRYVTGDCRSKAPVTELVCSGCCGPARLLPNAIGRGKMWPSGPDPRCIPDRYRAOR 120
DB 84 HYRFVTGDCRSKAPVTELVCSGCCGPARLLPNAIGRVKMWPNPDPFCIPDRYRAOR 143
QY 121 VOLLCGGAAPRAKRVLVASCCKRLTRFHNOSLKDFTGTEAARPOKGRKPRPARSAK 180
DB 144 VOLLCGGAAPRAKRVLVASCCKRLTRFHNOSLKDFTGTEAARPOKGRKPRPARSAK 203
QY 181 ANQAELENAY 190
DB 204 ANQAELENAY 213

RESULT 6
US-09-449-218D-12
Sequence 12, Application US/09449218D
Patent No. 6395511
GENERAL INFORMATION:
APPLICANT: Brunkow, Mary E.
APPLICANT: Galas, David J.
APPLICANT: Kovacevich, Brian
APPLICANT: Mulligan, John T.
APPLICANT: Paepert, Bryan W.
APPLICANT: Van Ness, Jeffrey
APPLICANT: Winkler, David G.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
FILE REFERENCE: 240083.508
CURRENT APPLICATION NUMBER: US/09/449,218D
CURRENT FILING DATE: 1999-11-24
NUMBER OF SEQ ID NOS: 45
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 12
LENGTH: 211
TYPE: PRT
ORGANISM: Mus musculus
US-09-449-218D-12

Query Match 90.7%; Score 951; DB 4; Length 211;
Best Local Similarity 89.5%; Pred. No. 3.4e-93;
Matches 170; Conservative 10; Mismatches 8; Indels 2; Gaps 1;

QY 1 QGWAQKNDATETIIPELGEYEPPEPELENNKTMRANENGRPPHPEPEKDVSEYSCREL 60
DB 24 QGWAQKNDATETIIPGLREYEPPEPELENNKTMRANENGRPPHPPYDQVSEYSCREL 81
QY 61 HFRYVTGDCRSKAPVTELVCSGCCGPARLLPNAIGRGKMWPSGPDPRCIPDRYRAOR 120
DB 82 HYRFVTGDCRSKAPVTELVCSGCCGPARLLPNAIGRVKMWPNPDPFCIPDRYRAOR 141
QY 121 VOLLCGGAAPRAKRVLVASCCKRLTRFHNOSLKDFTGTEAARPOKGRKPRPARSAK 180

DB 142 VOLLCGGAAPRAKRVLVASCCKRLTRFHNOSLKDFTGTEAARPOKGRKPRPARSAK 201
QY 181 ANQAELENAY 190
DB 202 ANQAELENAY 211

RESULT 7
US-09-449-218D-16
Sequence 16, Application US/09449218D
Patent No. 6395511
GENERAL INFORMATION:
APPLICANT: Brunkow, Mary E.
APPLICANT: Galas, David J.
APPLICANT: Kovacevich, Brian
APPLICANT: Mulligan, John T.
APPLICANT: Paepert, Bryan W.
APPLICANT: Van Ness, Jeffrey
APPLICANT: Winkler, David G.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
FILE REFERENCE: 240083.508
CURRENT APPLICATION NUMBER: US/09/449,218D
CURRENT FILING DATE: 1999-11-24
NUMBER OF SEQ ID NOS: 45
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 16
LENGTH: 176
TYPE: PRT
ORGANISM: Bos taurus
US-09-449-218D-16

Query Match 85.9%; Score 901.5; DB 4; Length 176;
Best Local Similarity 93.8%; Pred. No. 4.7e-88;
Matches 166; Conservative 2; Mismatches 8; Indels 1; Gaps 1;

QY 8 NDATETIIPELGEYEPPEPELENNKTMRANENGRPPHPEPEKDVSEYSCRELHFTRYVT 67
DB 1 NDATETIIPELGEYEPPEPELENNKTMRANENGRPPHPEPEKDVSEYSCRELHFTRYVT 59
QY 68 DPCRSKAPVTELVCSGCCGPARLLPNAIGRGKMWPSGPDPRCIPDRYRAORVOLLCPG 127
DB 60 DPCRSKAPVTELVCSGCCGPARLLPNAIGRGKMWPSGPDPRCIPDRYRAORVOLLCPG 119
QY 128 GAAPRAKRVLVASCCKRLTRFHNOSLKDFTGTEAARPOKGRKPRPARSAKANAQ 184
DB 120 GAAPRAKRVLVASCCKRLTRFHNOSLKDFTGTEAARPOKGRKPRPARSAKANAQ 176

RESULT 8
US-09-188-930-159
Sequence 159, Application US/09188930A
Patent No. 6150502
GENERAL INFORMATION:
APPLICANT: Watson, James D.
APPLICANT: Strachan, Lorna
APPLICANT: Sleeman, Matthew
APPLICANT: Ornust, Rene
APPLICANT: Murison, James Greg
TITLE OF INVENTION: Compositions and Methods Isolated From Skin Cells
FILE REFERENCE: 11000.1011C1
CURRENT APPLICATION NUMBER: US/09/188,930A
CURRENT FILING DATE: 1998-11-09
NUMBER OF SEQ ID NOS: 348
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 159
LENGTH: 206
TYPE: PRT
ORGANISM: mouse
US-09-188-930-159

Query Match 32.8%; Score 344; DB 4; Length 206;

MEDIUM TYPE: 3.5 INCH DISKETTE
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WORD PERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/468,847B
FILING DATE: 6 June 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: MULLINS, J.G.
REGISTRATION NUMBER: 33,073
REFERENCE/DOCKET NUMBER: 325800-442
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 206 AMINO ACIDS
TYPE: AMINO ACID
STRANDEDNESS:
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
US-08-468-847B-20

Query Match 32.5%; Score 341; DB 1; Length 206;
Best Local Similarity 42.7%; Pred. No. 1,5e-28;

Matches 79; Conservative 30; Mismatches 62; Indels 14; Gaps 8;

QY 5 AFNDATTEILPELGEYEP-PPELENNKTNRAENGGRPPHPEFTKDV---EYSCREL 60
DB 23 AFNDATTEIL--YSHVYKVPVPAHSSNSTLNQARNGR--HFSNTGLDENTRYQVACREL 78
QY 61 HFRYRYTDCGRSAKRYTELVCSCGCPARLLPNAIGRG---KWM-RPGCPDFRCIPDRY 116
DB 79 RSKYISDGGCTISPLKELVCAEGCLPLVLPNMWIGGYGTGYWRSRSEQWRNVNDXT 138
QY 117 RAQRVOLLGCGEAPRARKYRLVASCKKRLTFPHNSELDKDFGTAEAPQKQRKRPRA 176
DB 139 RTGRIOQCCDQ--STRYKLTIVTACKCKRYTRQHNSHNSFMSPAFVQHHREKKA 197
QY 177 -RSAX 180
DB 198 SKSSX 202

RESULT 12
US-09-449-218D-42
Sequence 42, Application US/09449218D
Patent No. 6395511

GENERAL INFORMATION:
APPLICANT: Brunkow, Mary E.
APPLICANT: Galas, David J.
APPLICANT: Kovacevich, Brian
APPLICANT: Mulligan, John T.
APPLICANT: Paepert, Bryan W.
APPLICANT: Van Ness, Jeffrey
APPLICANT: Winkler, David G.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
FILE REFERENCE: 240083.508
CURRENT APPLICATION NUMBER: US/09/449,218D
CURRENT FILING DATE: 1999-11-24
NUMBER OF SEQ ID NOS: 45
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 42
LENGTH: 184
TYPE: PRT
ORGANISM: Homo sapiens
US-09-449-218D-42

Query Match 8.5%; Score 89.5; DB 4; Length 184;
Best Local Similarity 24.1%; Pred. No. 0.06;
Matches 38; Conservative 22; Mismatches 57; Indels 41; Gaps 9;

QY 18 GEYEPPELENNK-----TNRAENGGRPPHPEFTKDVSEYSCRELHFT--R 64
DB 31 GAIP-PDKAQHNDSQTSPOQPSRNRGRGQGRGTMPGE--EVLSSQBALHVTBRK 87
QY 65 YVTDGPCR-----SAKPYTELVCSCGCPARLLPNAIGRKWMRPSGDPFR- 110
DB 88 YLRDMCKTQPLKQTHIEBGCNSRTIIRFCYGCN--SFYIPRIH-----RKIEGSFOS 140
QY 111 ---CIPDRYRAQVOLLCPGGEAP-RARKYRLVASCKK 144
DB 141 CSFCKPKKFTTMVTLNCPDLQPPTKKKRYRVKQCRC 178

RESULT 13
US-09-040-229B-2
Sequence 2, Application US/09040229B
Patent No. 6432410

GENERAL INFORMATION:
APPLICANT: Harland, Richard
Hsu, David
TITLE OF INVENTION: Morphogenic Proteins
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESS: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 75 DENISE DRIVE
CITY: HILLSBOROUGH
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94010

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/040,229B
FILING DATE: 13-Mar-1998
CLASSIFICATION: <unknown>

ATTORNEY/AGENT INFORMATION:
NAME: OSMAN, RICHARD A.
REGISTRATION NUMBER: 36,627

REFERENCE/DOCKET NUMBER: B97-020-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 343-4341
TELEFAX: (650) 343-4342

SEQUENCE CHARACTERISTICS:
SEQUENCE FOR SEQ ID NO: 2:
LENGTH: 184 amino acids
TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-040-229B-2

Query Match 8.5%; Score 89.5; DB 4; Length 184;
Best Local Similarity 24.1%; Pred. No. 0.06;
Matches 38; Conservative 22; Mismatches 57; Indels 41; Gaps 9;

QY 18 GEYEPPELENNK-----TNRAENGGRPPHPEFTKDVSEYSCRELHFT--R 64
DB 31 GAIP-PDKAQHNDSQTSPOQPSRNRGRGQGRGTMPGE--EVLSSQBALHVTBRK 87
QY 65 YVTDGPCR-----SAKPYTELVCSCGCPARLLPNAIGRKWMRPSGDPFR- 110
DB 88 YLRDMCKTQPLKQTHIEBGCNSRTIIRFCYGCN--SFYIPRIH-----RKIEGSFOS 140
QY 111 ---CIPDRYRAQVOLLCPGGEAP-RARKYRLVASCKK 144
DB 141 CSFCKPKKFTTMVTLNCPDLQPPTKKKRYRVKQCRC 178

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 28, 2003, 18:03:21 ; Search time 40.1204 Seconds
(without alignments)
4861.524 Million cell updates/sec

Title: US-09-867-274-3

Perfect score: 636
Sequence: 1 atgcagccctcactagcccc.....agctgagagagcctcactag 636

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA: *
1: /cgn2_6/prodata/1/ina/5A_COMB.seq: *
2: /cgn2_6/prodata/1/ina/5B_COMB.seq: *
3: /cgn2_6/prodata/1/ina/6A_COMB.seq: *
4: /cgn2_6/prodata/1/ina/6B_COMB.seq: *
5: /cgn2_6/prodata/1/ina/PCTUS_COMB.seq: *
6: /cgn2_6/prodata/1/ina/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	634.4	99.7	638	US-09-449-218D-11	Sequence 11, Appl
2	557.6	87.7	674	US-09-449-218D-13	Sequence 13, Appl
3	504.8	79.4	642	US-09-449-218D-9	Sequence 9, Appl
4	501.6	78.9	642	US-09-449-218D-45	Sequence 45, Appl
5	501.6	78.9	2301	US-09-449-218D-1	Sequence 1, Appl
6	501.6	78.9	2301	US-09-449-218D-5	Sequence 5, Appl
7	500	78.6	2301	US-09-449-218D-3	Sequence 3, Appl
8	500	78.6	2301	US-09-449-218D-7	Sequence 7, Appl
9	426.6	67.1	35828	US-09-449-218D-17	Sequence 17, Appl
10	415.8	65.4	532	US-09-449-218D-15	Sequence 15, Appl
11	367	57.7	9301	US-09-449-218D-18	Sequence 18, Appl
12	81.6	12.8	962	US-09-188-930-40	Sequence 40, Appl
13	81.6	12.8	962	US-09-188-930-209	Sequence 209, Appl
14	67.2	10.6	900	US-08-468-8478-1	Sequence 1, Appl
15	65.6	10.3	1692	US-09-602-877A-92	Sequence 92, Appl
16	52.6	8.3	71989	US-09-443-501A-2	Sequence 2, Appl
17	48.8	7.7	44377	US-08-804-227C-7	Sequence 7, Appl
18	48.8	7.7	44377	US-08-804-198-1	Sequence 1, Appl
19	44.4	7.0	8257	US-09-162-970B-65	Sequence 65, Appl
20	44.2	6.9	3930	US-09-467-946-2	Sequence 2, Appl
21	44.2	6.9	3930	US-09-467-946-2	Sequence 2, Appl
22	42.8	6.7	11220	US-09-105-537-32	Sequence 32, Appl
23	42.8	6.7	36778	US-09-105-537-5	Sequence 5, Appl
24	42.8	6.7	38506	US-09-320-878-19	Sequence 19, Appl
25	42.4	6.7	2040	US-08-031-538-10	Sequence 10, Appl
26	42.2	6.6	4050	US-09-543-084A-26	Sequence 26, Appl
27	42.2	6.6	4093	US-09-543-084A-28	Sequence 28, Appl

28	42.2	6.6	4101	US-09-543-084A-27	Sequence 27, Appl
29	41.8	6.6	974	US-09-449-218D-13	Sequence 13, Appl
30	41.4	6.5	953	US-09-452-239-45	Sequence 45, Appl
31	41.4	6.5	1820	US-08-173-508-7	Sequence 7, Appl
32	41.4	6.5	1821	US-08-265-310-7	Sequence 7, Appl
33	41.4	6.5	1821	US-08-951-742-7	Sequence 7, Appl
34	41	6.4	43280	US-08-804-227C-1	Sequence 1, Appl
35	40.8	6.4	15664	US-08-402-282-3	Sequence 3, Appl
36	40.8	6.4	15664	US-08-508-004-3	Sequence 3, Appl
37	40.8	6.4	15664	US-08-402-066-3	Sequence 3, Appl
38	40.8	6.4	15664	US-08-402-068-3	Sequence 3, Appl
39	40.6	6.4	4098	US-09-268-866-1	Sequence 1, Appl
40	40.4	6.4	20235	US-07-642-734C-3	Sequence 3, Appl
41	40.4	6.4	20235	US-08-439-009A-3	Sequence 3, Appl
42	40.4	6.4	23673	US-09-773-816-1	Sequence 1, Appl
43	40.2	6.3	1506	US-09-158-767-7	Sequence 7, Appl
44	40.2	6.3	1506	US-09-158-767-8	Sequence 8, Appl
45	40.2	6.3	1506	US-09-158-767-9	Sequence 9, Appl

ALIGNMENTS

```
RESULT 1
US-09-449-218D-11
; Sequence 11, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepert, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 638
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-449-218D-11
Query Match          99.7%; Score 634.4; DB 4; Length 638;
Best Local Similarity 99.8%; Pred. No. 4.5e-135;
Matches 635; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGCAGCCCTCCTAGCCCGGCTCTCATCTGCTTACTTGTGAGGCTCTTGTGTCT 60
    |||
DB 1 ATGCAGCCCTCCTAGCCCGGCTCTCATCTGCTTACTTGTGAGGCTCTTGTGTCT 60
QY 61 GTTGAAGGCGGAGGCGGAGGCTTCTGAGAAATGATGACAGAGGCTCTTCTTGTCT 120
    |||
DB 61 GTTGAAGGCGGAGGCGGAGGCTTCTGAGAAATGATGACAGAGGCTCTTCTTGTCT 120
QY 121 GAGAGTACCCCGAGGCTCTCTCTGAGAAATGATGACAGAGGCTCTTCTTGTCT 180
    |||
DB 121 GAGAGTACCCCGAGGCTCTCTCTGAGAAATGATGACAGAGGCTCTTCTTGTCT 180
QY 181 GGCAGACCTTCCCATCATCTCTTATGACGCAAGATGTGTGAGTGTGAGTGTGAGT 240
    |||
DB 181 GGCAGACCTTCCCATCATCTCTTATGACGCAAGATGTGTGAGTGTGAGTGTGAGT 240
QY 241 CTGCACTTACACCGGCTTCTGACAGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGG 300
    |||
DB 241 CTGCACTTACACCGGCTTCTGACAGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGG 300
QY 301 TTGTGTGTCTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 360
    |||
```

Db	301	TTGGTGTCTCTCGGCGCAATGGGGCCCGCGCGCTGTGCTGCCAAGCGCATCGGGCGGCTG	360
QY	361	AAGTGTGGCGCCCGGAACGGAACCGGATTTCCGTGCAATCCGGATGCTACCGCGCGCAG	420
Db	361	AAGTGTGGCGCCCGGAACGGAACCGGATTTCCGTGCAATCCGGATGCTACCGCGCGCAG	420
QY	421	CGGGTCGACGCTGTGTGTCCCGGGGGGCGCGCGCGCGCTTGTGCGCAAGTGTGCTTGTGTG	480
Db	421	CGGGTCGACGCTGTGTGTCCCGGGGGGCGCGCGCGCGCTTGTGCGCAAGTGTGCTTGTGTG	480
QY	481	GCTCTGTGCAAGTCAAGCGCGCTCAACCGCTTCCACAACCAAGTGTGAGACTCAAGAACTTC	540
Db	481	GCTCTGTGCAAGTCAAGCGCGCTCAACCGCTTCCACAACCAAGTGTGAGACTCAAGAACTTC	540
QY	541	GGGCGCGAGACCGCGCGCGCGCGCAAGAGGTTGCAAGCCCGCGCGCGCTCCGGGAGCC	600
Db	541	GGGCGCGAGACCGCGCGCGCGCGCAAGAGGTTGCAAGCCCGCGCGCGCTCCGGGAGCC	600
QY	601	AAAGCCAAACGAGCGGAGCTGTGGAAGCGCTCAATAG	636
Db	601	AAAGCCAAACGAGCGGAGCTGTGGAAGCGCTCAATAG	636

```

RESULT 2
US-09-449-218D-13
; Sequence 13. Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepel, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; TITLE OF INVENTION: BONE MINERALIZATION
; FILE REFERENCE: 240083 508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 13
; LENGTH: 674
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; US-09-449-218D-13

```

	Query Match	87.7%	Score 557.6;	DB 4;	Length 674;
	Best Local Similarity	93.0%;	Pred. No. 1e-117;		
	Matches 597;	Conservative 0;	Mismatches 39;	Indels 6;	Gaps 1
QY	1 ATGCAGCCCTCATTACGCCCGGCTCATCTGTGCTTATGGCAAGCTGCTTGTTGCT	60			
Db	33 ATGCAGCTTCATCATTAACCCCTTGTGCTTGCCTTGTTCATGAAGCCCTTGTTGCT	92			
QY	61 GTGAGAGGCACGAGGGTGTGCAAGCTTACAGAATGATGCAAGAGTTCATCCAGGGCTT	120			
Db	93 GTGAGAGGCCAGGGGTGTGCAAGCTTCAAAGATGATGCAAGAAATCATCTCCGGAGCTC	152			
QY	121 GGAGACTGATCCCGAAGCTTCTCT-----CCTGAGAACAAACGACCATTGAACCGGGCGAG	174			
Db	153 AGAGAGTACCCGAAGCTTCTTAGGAACTGAGAAACAACGACCATTGAACCGGGCGAG	212			
QY	175 AATGAGAGCGAGACTTCCCACATCCCTATGAGCGCAAGAATGTGTCGAGTACAGCTGC	234			
Db	213 AAGGAGGCGAGACCCCCCACATCTTTATGACACAAAGACGTGTCCAGTACAGCTGC	272			
QY	235 CGGAGCTGATCAACACCCGCTTCTTGACAGAGGCGCATGCGCGAGCGCAAGCCGGTCT	294			
Db	273 CGGAGCTGATCAACACCGCTTCTTGAGCCAGCGGCTGTCGCGAGTGCAGCCGGTCT	332			

Qy	295	ACCACTGTGTGTGCTCCGACCATGTGGGCGCGCGGCGCTGTCCCAAGCATCGGG	3.54
Db	333	ACCGAATTGTGTGTGTCCGGGCAATGTGGGCGCGCGGCTGTCCCAAGCATCGGG	3.92
Qy	355	CGCGTAAATGTGTGGCGGCGCAAGGAACTTCGCGTGATCCGGATTCGGTACCGC	41.4
Db	393	CGGTTAAATGTGTGGCGGCGCAAGGAACTTCGCGTGATCCGGATTCGGTACCGC	45.2
Qy	415	GCGCAGCGGGATCGACTCTGTGCCCCGGGGGCGCGCGCGCGTCTCGCGAAGTGTGT	47.4
Db	453	GCGCAACGGGTGTCACTCTGTGTGCCGGGGGCGCGCGCGCGTCTCGCGAAGTGTGT	51.2
Qy	475	CTGTGTGCTCTGTGCAAGTGTGACGCGCTTACCGCTTCCACAACCAAGTGTGAAGTCAAG	53.4
Db	513	CTGTGTGCTCTGTGCAAGTGTGACGCGCTTACCGCTTCCACAACCAAGTGTGAAGTCAAG	57.2
Qy	535	GACTTGGGGCCCGGAACCGCGCGGCGCGCAAAAGGTGTGAACCGCGCGCGCGCGCGCGG	59.4
Db	573	GACTTGGGACTGTGAACTCGCGCGGCGCGCAAAAGGTGTGAACCGCGCGCGCGCGCGCGG	63.2
Qy	595	GGAGCCAAAGCCAACTCAGCGGAGTGTGAAGCGCTTACTAG	63.6
Db	633	GGAGCCAAAGCCAACTCAGCGGAGTGTGAAGCGCTTACTAG	67.4

```

: RESULT 3
: US-09-449-218D-9
: Sequence 9, Application US/095449218BD
: Patent No. 6395511
: GENERAL INFORMATION:
: APPLICANT: Brunkow, Mary E.
: APPLICANT: Galas, David J.
: APPLICANT: Kovacevich, Brian
: APPLICANT: Mulligan, John T.
: APPLICANT: Peaper, Bryan W.
: APPLICANT: Van Ness, Jeffrey
: APPLICANT: Winkler, David G.
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
: FILE REFERENCE: BONE MINERALIZATION
: CURRENT APPLICATION NUMBER: US/09/449,218D
: CURRENT FILING DATE: 1999-11-24
: NUMBER OF SEQ ID NOS: 45
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 9
: LENGTH: 642
: TYPE: DNA
: ORGANISM: Cercopithecus pygerythrus
: US-09-449-218D-9

```

Query Match	Score	DB #	Length
79.4%	504.8	DB 4	642
Best Local Similarity	87.9%	Pred. No. 8.6e-106	
Matches	554	Conservative	0
		Mismatches	72
		Indels	6
		Gaps	1

QY 295 ACCGAGTTGCTGCTCCGCGCAGTGCAGCCCGCGGCTGTGCCACGCCATCGG 354
DB 301 ACCGAGTTGCTGCTCCGCGCAGTGCAGCCCGCGGCTGTGCCACGCCATCGG 360
QY 355 CGCGTGAAGTGTGGCGCGCGGAAAGCAACGGATTTCGCTGATCCCGATCGCTACCG 414
DB 361 CGCGGCAAGTGTGGCGCGCGGAAAGCAACGGATTTCGCTGATCCCGATCGCTACCG 420
QY 415 GCGCAGGAGTGTGAGCTGTGTGCTCCCGGCGCGCGGCTGTGCCAGTGTGCT 474
DB 421 GCGCAGGAGTGTGAGCTGTGTGCTCCCGGCGCGCGGCTGTGCCAGTGTGCT 480
QY 475 CTGTGTGCTGTGTGCAAGTGTGCAAGCTGTGCAAGCTGTGCAAGCTGTGCAAG 534
DB 481 CTGTGTGCTGTGTGCAAGTGTGCAAGCTGTGCAAGCTGTGCAAGCTGTGCAAG 540
QY 535 GACTTGGGCGCGGAGACCGCGCGCGCGGAGAGGTGTGCAAGCGCGCGCGCGG 594
DB 541 GACTTGGGCGCGGAGACCGCGCGCGCGGAGAGGTGTGCAAGCGCGCGCGCGG 600
QY 595 GGAGCCAAAGCCAAACGAGCGGAGCTGTGAGAAAGCTGTACTAG 636
DB 601 GGAGCCAAAGCCAAACGAGCGGAGCTGTGAGAAAGCTGTACTAG 642

RESULT 4
US-09-449-218D-45
Sequence 45, Application US/09449218D
Patent No. 6395511
GENERAL INFORMATION:
APPLICANT: Brunkow, Mary E.
APPLICANT: Galas, David J.
APPLICANT: Kovacevich, Brian
APPLICANT: Mulligan, John T.
APPLICANT: Paepker, Bryan W.
APPLICANT: Van Ness, Jeffrey
APPLICANT: Winkler, David G.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
TITLE OF INVENTION: BONE MINERALIZATION
FILE REFERENCE: 240083.508
CURRENT APPLICATION NUMBER: US/09/449,218D
CURRENT FILING DATE: 1999-11-24
NUMBER OF SEQ ID NOS: 45
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 45
LENGTH: 642
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(639)
US-09-449-218D-45

Query Match 78.9%; Score 501.6; DB 4; Length 642;
Best Local Similarity 87.5%; Pred. No. 4.5e-105;
Matches 562; Conservative 0; Mismatches 74; Indels 6; Gaps 1;

QY 1 ATGACGCTCTCACTAGCCCGCTGTCTATCTGCTTGTGACGCTGTCTGTCT 60
DB 1 ATGACGCTCTCACTAGCCCGCTGTCTATCTGCTTGTGACGCTGTCTGTCT 60
QY 61 GTGAGAGGCGGAGGAGGTGTGCAAGCTGTGAGAAATGATGCAAGGCAATCCGAGGCTT 120
DB 61 GTGAGAGGCGGAGGAGGTGTGCAAGCTGTGAGAAATGATGCAAGGCAATCCGAGGCTT 120
QY 121 GGAGAGTACCCCGAGCTCTCTCC-----TGAGAACAAACGAGCAATGAAACCGGCGGAG 174
DB 121 GGAGAGTACCCCGAGCTCTCTCC-----TGAGAACAAACGAGCAATGAAACCGGCGGAG 180
QY 175 AATGAGAGCAAGCTCTCCCAACATCTCTATGAGCGCAAGAGTGTGTCCAGTACAGCTGC 234
DB 181 AATGAGAGCAAGCTCTCCCAACATCTCTATGAGCGCAAGAGTGTGTCCAGTACAGCTGC 240

QY 225 CGCGAGTGTGCTGTGCAAGCTGTGCTGTGCAAGCTGTGCAAGCTGTGCAAG 294
DB 241 CGCGAGTGTGCTGTGCAAGCTGTGCTGTGCAAGCTGTGCAAGCTGTGCAAG 300
QY 295 ACCGAGTTGCTGCTCCGCGCAGTGCAGCCCGCGGCTGTGCCACGCCATCGG 354
DB 301 ACCGAGTTGCTGCTCCGCGCAGTGCAGCCCGCGGCTGTGCCACGCCATCGG 360
QY 355 CGCGTGAAGTGTGGCGCGCGGAAAGCAACGGATTTCGCTGATCCCGATCGCTACCG 414
DB 361 CGCGGCAAGTGTGGCGCGCGGAAAGCAACGGATTTCGCTGATCCCGATCGCTACCG 420
QY 415 GCGCAGGAGTGTGAGCTGTGTGCTCCCGGCGCGCGGCTGTGCCAGTGTGCT 474
DB 421 GCGCAGGAGTGTGAGCTGTGTGCTCCCGGCGCGCGGCTGTGCCAGTGTGCT 480
QY 475 CTGTGTGCTGTGTGCAAGTGTGCAAGCTGTGCAAGCTGTGCAAGCTGTGCAAG 534
DB 481 CTGTGTGCTGTGTGCAAGTGTGCAAGCTGTGCAAGCTGTGCAAGCTGTGCAAG 540
QY 535 GACTTGGGCGCGGAGACCGCGCGCGCGGAGAGGTGTGCAAGCGCGCGCGCGG 594
DB 541 GACTTGGGCGCGGAGACCGCGCGCGCGGAGAGGTGTGCAAGCGCGCGCGCGG 600
QY 595 GGAGCCAAAGCCAAACGAGCGGAGCTGTGAGAAAGCTGTACTAG 636
DB 601 GGAGCCAAAGCCAAACGAGCGGAGCTGTGAGAAAGCTGTACTAG 642

RESULT 5
US-09-449-218D-1
Sequence 1, Application US/09449218D
Patent No. 6395511
GENERAL INFORMATION:
APPLICANT: Brunkow, Mary E.
APPLICANT: Galas, David J.
APPLICANT: Kovacevich, Brian
APPLICANT: Mulligan, John T.
APPLICANT: Paepker, Bryan W.
APPLICANT: Van Ness, Jeffrey
APPLICANT: Winkler, David G.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
TITLE OF INVENTION: BONE MINERALIZATION
FILE REFERENCE: 240083.508
CURRENT APPLICATION NUMBER: US/09/449,218D
CURRENT FILING DATE: 1999-11-24
NUMBER OF SEQ ID NOS: 45
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1
LENGTH: 2301
TYPE: DNA
ORGANISM: Homo sapien
US-09-449-218D-1

Query Match 78.9%; Score 501.6; DB 4; Length 2301;
Best Local Similarity 87.5%; Pred. No. 5.3e-105;
Matches 562; Conservative 0; Mismatches 74; Indels 6; Gaps 1;

QY 1 ATGACGCTCTCACTAGCCCGCTGTCTATCTGCTTGTGACGCTGTCTGTCT 60
DB 48 ATGACGCTCTCACTAGCCCGCTGTCTATCTGCTTGTGACGCTGTCTGTCT 107
QY 61 GTGAGAGGCGGAGGAGGTGTGCAAGCTGTGAGAAATGATGCAAGGCAATCCGAGGCTT 120
DB 108 GTGAGAGGCGGAGGAGGTGTGCAAGCTGTGAGAAATGATGCAAGGCAATCCGAGGCTT 167
QY 121 GGAGAGTACCCCGAGCTCTCTCC-----TGAGAACAAACGAGCAATGAAACCGGCGGAG 174
DB 168 GGAGAGTACCCCGAGCTCTCTCC-----TGAGAACAAACGAGCAATGAAACCGGCGGAG 227
QY 175 AATGAGAGCAAGCTCTCCCAACATCTCTATGAGCGCAAGAGTGTGTCCAGTACAGCTGC 234

Db 228 AACGAGGCGCGCTCCCAACCCCTTTGAGACAAAGACGTGTCCAGTACAGCTGC 287
Qy 235 CGGAGAGTGAATAACCCGCTTCTGACAGACGCCCAAGCCGACGCGGCTC 294
Db 288 CGCAGAGTGAATAACCCGCTTCTGACAGACGCCCAAGCCGAGCTC 347
Qy 295 ACCGAGTGTGTCTCCGAGCAGTGCAGCGCCCGGCGCTGTGACCAAGCCATCGAG 354
Db 348 ACCGAGTGTGTCTCCGAGCAGTGCAGCGCCCGGCGCTGTGACCAAGCCATCGAG 407
Qy 355 CGCGTGAAGTGTGTCTCCGAGCAGTGCAGCGCCCGGCGCTGTGACCAAGCCATCGAG 414
Db 408 CGCGTGAAGTGTGTCTCCGAGCAGTGCAGCGCCCGGCGCTGTGACCAAGCCATCGAG 467
Qy 415 GCGAGCGGAGTGTGTCTCCGAGCAGTGCAGCGCCCGGCGCTGTGACCAAGCCATCGAG 474
Db 468 GCGAGCGGAGTGTGTCTCCGAGCAGTGCAGCGCCCGGCGCTGTGACCAAGCCATCGAG 527
Qy 475 CTGTGAGCTGTGTGAGTGAAGCGCTCAACCGCTTCCCAACCAAGTGGAGCTCAAG 534
Db 528 CTGTGAGCTGTGTGAGTGAAGCGCTCAACCGCTTCCCAACCAAGTGGAGCTCAAG 587
Qy 535 GACTTCGAGCGGAGACCGCGCGCGCGCGAGAGGTGCGCAAGCGCGCGCGCGCGCG 594
Db 588 GACTTCGAGCGGAGACCGCGCGCGCGCGAGAGGTGCGCAAGCGCGCGCGCGCGCG 647
Qy 595 GAGACCAAGCCAAACCGAGCGGAGCTGAGAAACCGCTACTAG 636
Db 648 AGCGCCAAAGCCAAACCGAGCGGAGCTGAGAAACCGCTACTAG 689

RESULT 6

US-09-449-218D-5
; Sequence 5, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepfer, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; TITLE OF INVENTION: BONE MINERALIZATION
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 2301
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-449-218D-5

Query Match 78.9%; Score 501.6; DB 4; Length 2301;
Best Local Similarity 87.5%; Pred. No. 5.3e-105;
Matches 562; Conservative 0; Mismatches 74; Indels 6; Gaps 1;
Qy 1 ATGAGCCCTCACTAGCCCGGCTGCTCATGCTACTTGTGACGCGCTTGTGCT 60
Db 48 ATGAGCTCCCACTGCGCTGTGTCTCATGCTGTGTGACAGAGCTTCCGCTA 107
Qy 61 GTGAGGCGCAGGGGTGCGAAGCTTCAAGATGATGCCAAGAGTCTATCCAGGGCTT 120
Db 108 GTGAGGCGCAGGGGTGCGAAGCTTCAAGATGATGCCAAGAGTCTATCCAGGGCTT 167
Qy 121 GGAAGTACCCCGAGCTCTCTCC-----TGAGAACAAACGAGACCAAGACCGGGCGAG 174
Db 168 GGAAGTACCCCGAGCTCTCTCC-----TGAGAACAAACGAGACCAAGACCGGGCGAG 227
Qy 175 AATGAGGCGAGACTCTCCCAATCTCTATGACGCCAAAGATGTGTCCAGTACAGCTGC 234

Db 228 AACGAGGCGCGCTCCCAACCCCTTTGAGACAAAGACGTGTCCAGTACAGCTGC 287
Qy 235 CGGAGAGTGAATAACCCGCTTCTGACAGACGCCCAAGCCGACGCGGCTC 294
Db 288 CGCAGAGTGAATAACCCGCTTCTGACAGACGCCCAAGCCGAGCTC 347
Qy 295 ACCGAGTGTGTCTCCGAGCAGTGCAGCGCCCGGCGCTGTGACCAAGCCATCGAG 354
Db 348 ACCGAGTGTGTCTCCGAGCAGTGCAGCGCCCGGCGCTGTGACCAAGCCATCGAG 407
Qy 355 CGCGTGAAGTGTGTCTCCGAGCAGTGCAGCGCCCGGCGCTGTGACCAAGCCATCGAG 414
Db 408 CGCGTGAAGTGTGTCTCCGAGCAGTGCAGCGCCCGGCGCTGTGACCAAGCCATCGAG 467
Qy 415 GCGAGCGGAGTGTGTCTCCGAGCAGTGCAGCGCCCGGCGCTGTGACCAAGCCATCGAG 474
Db 468 GCGAGCGGAGTGTGTCTCCGAGCAGTGCAGCGCCCGGCGCTGTGACCAAGCCATCGAG 527
Qy 475 CTGTGAGCTGTGTGAGTGAAGCGCTCAACCGCTTCCCAACCAAGTGGAGCTCAAG 534
Db 528 CTGTGAGCTGTGTGAGTGAAGCGCTCAACCGCTTCCCAACCAAGTGGAGCTCAAG 587
Qy 535 GACTTCGAGCGGAGACCGCGCGCGCGCGAGAGGTGCGCAAGCGCGCGCGCGCG 594
Db 588 GACTTCGAGCGGAGACCGCGCGCGCGCGAGAGGTGCGCAAGCGCGCGCGCGCGCG 647
Qy 595 GAGACCAAGCCAAACCGAGCGGAGCTGAGAAACCGCTACTAG 636
Db 648 AGCGCCAAAGCCAAACCGAGCGGAGCTGAGAAACCGCTACTAG 689

RESULT 7

US-09-449-218D-3
; Sequence 3, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepfer, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; TITLE OF INVENTION: BONE MINERALIZATION
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 2301
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-449-218D-3

Query Match 78.6%; Score 500; DB 4; Length 2301;
Best Local Similarity 87.4%; Pred. No. 1.2e-104;
Matches 561; Conservative 0; Mismatches 75; Indels 6; Gaps 1;
Qy 1 ATGAGCCCTCACTAGCCCGGCTGCTCATGCTACTTGTGACGCGCTTGTGCT 60
Db 48 ATGAGCTCCCACTGCGCTGTGTCTCATGCTGTGTGACAGAGCTTCCGCTA 107
Qy 61 GTGAGGCGCAGGGGTGCGAAGCTTCAAGATGATGCCAAGAGTCTATCCAGGGCTT 120
Db 108 GTGAGGCGCAGGGGTGCGAAGCTTCAAGATGATGCCAAGAGTCTATCCAGGGCTT 167
Qy 121 GGAAGTACCCCGAGCTCTCTCC-----TGAGAACAAACGAGACCAAGACCGGGCGAG 174
Db 168 GGAAGTACCCCGAGCTCTCTCC-----TGAGAACAAACGAGACCAAGACCGGGCGAG 227

QY	175	AATGAGGCAACACCTCCCAACATCCCTATAGACGCAAAAGATGTGTCCTCCAGTACAGCTGC	23.4
Db	228	AACGGAGGGCGGGCTTCCCAACACCCCTTTAGACCAAAAGACGTGTCTCAAGTACAGCTGC	28.7
QY	235	CGCAGACTGCATCAACCCCGTTCTTGA CAGACGGACCATGCGCAGCGCCCAACCGCGTTC	29.4
Db	288	CGCAGAGTGCACCTTCAACCCGCTTAGCGATGGGCGGTGCGCGACGCCCAACCGCGTTC	34.7
QY	295	AACCAATTGTGTGTCTCCGGGCAATGGCGGGCCCGCGCGGCTGCTGGCCCAACGCCATGGG	35.4
Db	348	ACCGAGGCTGTGTCTCCGGGCAATGGCGGGCCCGCGCGGCTGCTGGCCCAACGCCATGGG	40.7
QY	355	CGCGTGAAGTGTGGCGGCCCAACGAGATTTCCGCTGATCCCGATTCGCTTACCGC	41.4
Db	408	CGCGGCAAGTGTGTGGCGGCACTAAGTGTGGCGGCACTTCCGCTGATCCCGATTCGCTTACCGC	46.7
QY	415	CGCGAGCGGGTGCAGCTGTGTGCCCCGGGGGCGCGCGCGCGCTGCAGCAAGTGTGCT	47.4
Db	468	CGCGAGCGCGTGCAGCTGTGTGTGCCGTGTGTGAAGGGCGCGCGCGCGCGCAAGTGTGCG	52.7
QY	475	CTGGTGGCTGTGTGCACAGTGCAGAGGCGCTTCAACCGCTTCCCAACACAGTGCAGCTTCAAG	53.4
Db	528	CTGGTGGCTGTGTGCAGATGCAGAGGCGCTTCAACCGCTTCCCAACACAGTGCAGCTTCAAG	58.7
QY	535	GACTTTCGGGCGCGAGACCGCGCGGCGCGGAGAGGGTGTGCAGACCGCGGAGCTCCGGCGCCGG	59.4
Db	588	GACTTTCGGGAGCGAGAGCGCTTCGGCGCGGAGAGGGTGTGCAGAGCGCGGAGCTCCGGCGCCGG	64.7
QY	595	GGAGGCAAGGCAACCGAGCGGAGAGCTGTGAGAACCCCTTACTAG	63.6
Db	648	AGCGCAAGGCAACCGAGCGGAGAGCTGTGAGAACCGCTTACTAG	68.9

```

RESULT 8
US-09-449-218D-7
; Sequence 7, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepert, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; TITLE OF INVENTION: BONE MINERALIZATION
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 2301
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-449-218D-7

```

Query Match	78.6%	Score 500	DB 4	Length 2301
Best Local Similarity	87.4%	Pred. No. 1,2e-104		
Matches 561	Conservative 0	Mismatches 75	Indels 6	Gaps 1
Qy	1	ATGCAGCCTCACTAGCCCCCGTCCCTCATCTGCTACTTGTGACGCGTCTCTTGTGCT	60	
Db	48	ATGCAGCTCCCACTGACCCTGTGTCTCGTCTGCTCTGTACACACACGCTTCGTGTA	107	
Qy	61	GTGGAGGGCCAGGGGTGGGCAAGCTTCAAGAAATGATGCACAGAGGTCAATCCAGGGCTT	120	
Db	108	GTGGAGGGCCAGGGGTGGGCAAGCGCTTCAAGAAATGATGCACAGGAAATCAATCCGCGAGCTC	167	
Qy	121	GGAGAGTACCCGAGCGCTCTCC-----TGAGAACACAGACCATGAAACGGGGCGGAG	174	
Db	168	GGAGAGTACCCGAGCGCTTCAACCGAGCTGGAGAACACAGACCATGAAACCGGGCGGAG	227	

OY	175	AAAGAGGACACACTCCCAACATCCCTATGAGCGCAAAATATGTCCAGTACAGTGC	234
Db	228	AACGAGAGGCGGCTCTCCCAACACCCCTTTGAGACAAAGACGTGTCCAGTACAGTGC	287
OY	235	CGCAGCTGCACTACCCCGCTTCTGACAGACGGCCATGCGCAGACGCCAAGCCGGTC	294
Db	288	CGCAGCTGCACTTCAACCGCTTACGTACCGATGGGCGGTGCGCAGCGCCAAAGCCGGTC	347
OY	295	ACCGAGTTGGTGTCTCCGCGCAATGGCGGCCCGCGCGGCTGCTGCCAACGCCATTCGGG	354
Db	348	ACCGAGCTGGTGTCTCCGCGCAATGGCGGCCCGCGCGGCTGCTGCCAACGCCATTCGGC	407
OY	355	CGCGTGAAGTGTGTGGCGGCCGAACGGACCGGATTTTCGCTGCATCCCGGATCGCTACCGC	414
Db	408	CGCGGCAAGTGTGTGGCCACCTAATGGGGCCGACCTTCGCTGCATCCCGAACCGGTACCGC	467
OY	415	GCGCAGCGGGGTGCAGCTGTGTGCCCCGGGGGGCGCGCGCCGCGCTGTGCGCAAGTGTGCT	474
Db	468	GCGCGCGCCGTGTGCAGCTGTGTGTGCCGGTGTGTAGCGCGCGCGCGCCCAAGTGTGCGC	527
OY	475	CTGGTGGCTCTGTGTACAGTGCAGACGGCCTACCCGCTTTCACCAACAGTGCAGACTCAAG	534
Db	528	CTGGTGGCTCTGTGTCAATGTGCACACGGCCTTCAACCGCTTTCACCAACAGTGTGAGACTCAAG	587
OY	535	GACTTTCGGGACCGAGACCGCGCGGCGCCGACGAGAGGTTGCACAGCCCGCGGCGCGCCCGG	594
Db	588	GACTTTCGGGACCGAGAGCGCGCTCGCGCCGACAGAGGCGCGGAGAGCGCGCGCGCGCGCGG	647
OY	595	GGAGCCAAAGCCAAACGAGGGCGAGACTGTGAGAGACGCTTACTAG	636
Db	648	AGCGCCAAAGCCAAACGAGGGCGAGACTGTGAGAGACGCTTACTAG	689

```

RESULT 9
US-09-449-218D-17
; Sequence 17, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galias, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepfer, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; TITLE OF INVENTION: BONE MINERALIZATION
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 17
; LENGTH: 35828
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1) _ (35828)
; OTHER INFORMATION: n = A,T,C or G
US-09-449-218D-17

```

	Query Match	Similarity	Score	DB 4:	Length
Best Local	Similarity	95.8%	Pred. No. 6.5e-88		
Matches	438	Conservative	0	Mismatches	19
				Indels	0
				Gaps	0
QY	180	AGCGAGACCTCCCAACCAATCCCTATGACGCCAAGATGTCGAGTACAGTCGCCGA	239		
DB	20988	AGGTCGACAGCCCTCAGCGATCCCTCTCCGCAATGTGTCCGATACAGTCGCCGA	21047		
QY	240	GCTGCATACACCCGCTTCTTACACAGAGGCCCAATGCCGACGCCCAACCGGTCACCGA	299		

Db	21048	GGGACACTACACCGCGTCTCTBACAGACGGGCCCATATCGCAGACGGCAACGGCGTACCGCA	21107
Qy	300	GTTGGTGTCCTCCGCGCACTGCGGGCCCGCGCGGCTGCTGCCAAGCCATCGGCGCGT	359
Db	21108	GTTGGTGTCCTCCGCGCACTGCGGGCCCGCGCGGCTGCTGCCAAGCCATCGGCGCGT	21167
Qy	360	GAAAGTGATGGCGCCCGAAGCGAATTCGCTGTCATCCCGGATCGGTATACCGGGCGCA	419
Db	21168	GAAAGTGATGGCGCCCGAAGCGAATTCGCTGTCATCCCGGATCGGTATACCGGGCGCA	21227
Qy	420	GCGGGTGCACTGCTGTGCGCCCGGGGGCGCGCGCGCTGCGCAAGAGTGCTGTGGT	479
Db	21228	GCGGGTGCACTGCTGTGCGCCCGGGGGCGCGCGCGCTGCGCAAGAGTGCTGTGGT	21287
Qy	480	GCGCTGTGTCAAAGTGCAGACGGCTTCAACCGCTTCCAAACAGTCGGAAGTTCAAGAACTT	539
Db	21288	GCGCTGTGTCAAAGTGCAGACGGCTTCAACCGCTTCCAAACAGTCGGAAGTTCAAGAACTT	21347
Qy	540	CGGGCCCGAAGACCGCGCGCGCGCGCGCAAGAGGTGCGCAAGCGCGCGCGCGCGCGCGAGC	599
Db	21348	CGGGCCCGAAGACCGCGCGCGCGCGCGCAAGAGGTGCGCAAGCGCGCGCGCGCGCGAGC	21407
Qy	600	CAAGGCAACCGAGCGAGCTGAGAAAGCCTACTAG	636
Db	21408	CAAGGCAACCGAGCGAGCTGAGAAAGCCTACTAG	21444

```

RESULT 10
US-09-449-218D-15
; Sequence 15, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepfer, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; TITLE OF INVENTION: BONE MINERALIZATION
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1995-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 15
; LENGTH: 532
; TYPE: DNA
; ORGANISM: Bos taurus
US-09-449-218D-15

```

Query Match	Score	Length
65.4%; Best Local Similarity	415.8; 85.4%; Pred. No. 1.1e-85;	532;
Matches 466; Conservative	0; Mismatches 62; Indels 3; Gaps	

[illegible]

```

RESULT 11
US-09-449-218D-18
; Sequence 18, Application US/09449218D
; Patent No. 639511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepfer, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; TITLE OF INVENTION: BONE MINERALIZATION
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 18
; LENGTH: 9301
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-449-218D-18

```

	Query Match	Similarity	Score	DB	Length
	Best Local	86.2%	Pred. No. 1.6e-74		
	Matches 406;	Conservative	0;	Mismatches	65; Indels 0; Gaps 0
Qy	166	CGGCGGAGATGAGAGCGACCTTCCGACATCCCTATGACGCAAGAATGTCGAG	225		
Dp	3138	CAGGGGTGGCAGAGCGGGGCGCCACCTTCAGCGCGGCTCTTCACAGACGATGCCAG	3197		
Qy	226	TACAGCTGCCGCGAGCTGCATCAACCCGCTTCTGACAGCGGCCATGCGGAGGCC	285		
Dp	3198	TACAGCTGCCGCGAGCTGCATCTTACCCGCTACGACGACCGATGGGCGGTCCGACAGCGCC	3257		
Qy	286	AAGCGGTCAACCGATGTTGTGTGCTCCGGCCAGATGCGGCCCGCGGCGGCTGGCCCAAC	345		
Dp	3258	AAGCGGTCAACCGATGTTGTGTGCTCCGGCCAGATGCGGCCCGCGGCGGCTGGCCCAAC	3317		
Qy	346	GCCATCGGGGCGCGTAGATGATGGCGCCGAACGGACCGGATTTCCGCTCATTCGCGAT	405		
Dp	3318	GCCATCGGGGCGCGTAGATGATGGCGCATCTATGTGGGCCCACTTCCGCTCATTCGCCGAC	3377		
Qy	406	CGCTACCGCGCGAGGGGTGCAGTGTGTGCTCCCGGGGGCGCGGCGCTGCGCG	465		
Dp	3378	CGCTACCGCGCGAGGGGTGCAGTGTGTGCTCCCGGTGTGAGGCGCGCGCGCGCG	3437		
Qy	466	AAGGTGCTGTGTGGCTCTGTGCAAGTGCAGAGGCTCACCGCTTCCACCAACCAATGG	525		
Dp	3438	AAGGTGCTGTGTGGCTCTGTGCAAGTGCAGAGGCTCACCGCTTCCACCAACCAATGG	3497		

QY 526 GAGCTCAAGAGCTTGGGGCGGAGACCGCGCGGCGGAGAGGGGTGCTAGCCGCGGCGC 585
Db 3498 GAGCTCAAGAGCTTGGGGCGGAGACCGCGCGGCGGAGAGGGGTGCTAGCCGCGGCGC 3557
QY 586 GGGCGCGCGGAGCGCAAGCGCAACGAGCGGAGCTGAGAGACGCTTACG 636
Db 3558 CGCGCGCGGAGCGCGCAAGCGCAACGAGCGGAGCTGAGAGACGCTTACG 3608

RESULT 12
US-09-188-930-40
; Sequence 40, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Muriel, James Greg
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; TITLE OF INVENTION: and Methods For Their Use
; FILE REFERENCE: 11000.1011C1
; CURRENT APPLICATION NUMBER: US/09/188,930A
; CURRENT FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 40
; LENGTH: 962
; TYPE: DNA
; ORGANISM: mouse
US-09-188-930-40

Query Match 12.8%; Score 81.6; DB 3; Length 962;
Best Local Similarity 52.6%; Pred. No. 3.5e-10;
Matches 283; Conservative 0; Mismatches 234; Indels 21; Gaps 4;

QY 82 GCCTTCAGAGATGATGACCAAGAGCTATCCAGGGCTTGGAGAGTACCC---CGAGCCT 138
Db 113 GCTTTTAAAGATGATGACCAAGAGCTATCCAGGGCTTGGAGAGTACCC---CGAGCCT 172
QY 139 CCTCTGAGAACCAACGAGCCATGAAACCGGGCGGAGAGTGAAGGACGCTCC---CAC 195
Db 173 CACCCAGAGCAACAGACGACCTGATCAAGCAGAGATGAGAGGAGCATTTACATGAGC 232
QY 196 CATCCCTATGAGCCCAAGATGTCGAGTACAGCTGCGGAGCTGACATACACCGCG 255
Db 233 ACTGAGCTGATGCAAGAGTGAAGTCAAGTGGGCTGAGAGGAGCTGACCAACAA 292
QY 256 TTCTGACAGACGCGCCATGCGGAGCGCAAGCGGTCACCGAGTTGTGTCTCCGCG 315
Db 293 TACATTTTGGAGGCGGCAATGACCAAGCATCAAGCCCTCTGAGAGAGCTGTGTGCGCGG 352
QY 316 CAGTGGCGCGCGCGGCTGTGCTCCCAAGCGCATCGGCGCG---TGAAAG 363
Db 353 GAGTCTTGGCGCGCGGCTGTGCTCCCAAGCTGATCGAGAGGAGCTGACCAACAAAGTAC 412
QY 364 TGTGTGCGCGCGGAGAGCGGATTTCCGCTGATCCCGGATCGCTACCGCGCGAGCGG 423
Db 413 TGGAGCGCGGAGAGCTCTTGAAGAGTGGCGGTGTGAACGAGAGAGCGGACCGAGAGG 472
QY 424 GTGACAGCTGTGCGCGCGGAGCGGCGCGCTGCGGAGAGTGTGTGTGTGCGC 483
Db 473 ATCCAGCTGTGAGTGTGAAGAGCGG---AGCAGCGGACCTTCAAAATCAACCGTGTGTCG 529
QY 484 TGTGCAAGTGAAGCGCTTACCGGCTTCAACAAACAGTGAAGTGAAGTGAAGTTCGCGG 543
Db 530 GCGTCAAGTGAAGAGGTGACCCGTCAGCAACAGATTCAGGCAACATTTTGAAGC 589
QY 544 CCGAGAGCGCGCGCGGCGGAGAGGTGCAAGCGCGCGCGCGCGCGCGGAGAGCA 601
Db 590 GTGTGCGCGAGAGCG 647

RESULT 13
US-09-188-930-209
; Sequence 209, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Muriel, James Greg
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; TITLE OF INVENTION: and Methods For Their Use
; FILE REFERENCE: 11000.1011C1
; CURRENT APPLICATION NUMBER: US/09/188,930A
; CURRENT FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 209
; LENGTH: 962
; TYPE: DNA
; ORGANISM: Mouse
US-09-188-930-209

Query Match 12.8%; Score 81.6; DB 3; Length 962;
Best Local Similarity 52.6%; Pred. No. 3.5e-10;
Matches 283; Conservative 0; Mismatches 234; Indels 21; Gaps 4;

QY 82 GCCTTCAGAGATGATGACCAAGAGCTATCCAGGGCTTGGAGAGTACCC---CGAGCCT 138
Db 113 GCTTTTAAAGATGATGACCAAGAGCTATCCAGGGCTTGGAGAGTACCC---CGAGCCT 172
QY 139 CCTCTGAGAACCAACGAGCCATGAAACCGGGCGGAGAGTGAAGGACGCTCC---CAC 195
Db 173 CACCCAGAGCAACAGACGACCTGATCAAGCAGAGATGAGAGGAGCATTTACATGAGC 232
QY 196 CATCCCTATGAGCCCAAGATGTCGAGTACAGCTGCGGAGCTGACATACACCGCG 255
Db 233 ACTGAGCTGATGCAAGAGTGAAGTCAAGTGGGCTGAGAGGAGCTGACCAACAAAGTAC 292
QY 256 TTCTGACAGACGCGCCATGCGGAGCGCAAGCGGTCACCGAGTTGTGTCTCCGCG 315
Db 293 TACATTTTGGAGGCGGCAATGACCAAGCATCAAGCCCTCTGAGAGAGCTGTGTGCGCGG 352
QY 316 CAGTGGCGCGCGCGGCTGTGCTCCCAAGCGCATCGGCGCG---TGAAAG 363
Db 353 GAGTCTTGGCGCGCGGCTGTGCTCCCAAGCTGATCGAGAGGAGCTGACCAACAAAGTAC 412
QY 364 TGTGTGCGCGCGGAGAGCGGATTTCCGCTGATCCCGGATCGCTACCGCGCGAGCGG 423
Db 413 TGGAGCGCGGAGAGCTCTTGAAGAGTGGCGGTGTGAACGAGAGAGCGGACCGAGAGG 472
QY 424 GTGACAGCTGTGCGCGCGGAGCGGCGCGCTGCGGAGAGTGTGTGTGTGCGC 483
Db 473 ATCCAGCTGTGAGTGTGAAGAGCGG---AGCAGCGGACCTTCAAAATCAACCGTGTGTCG 529
QY 484 TGTGCAAGTGAAGCGCTTACCGGCTTCAACAAACAGTGAAGTGAAGTGAAGTTCGCGG 543
Db 530 GCGTCAAGTGAAGAGGTGACCCGTCAGCAACAGATTCAGGCAACATTTTGAAGC 589
QY 544 CCGAGAGCGCGCGCGGCGGAGAGGTGCAAGCGCGCGCGCGCGCGGAGAGCA 601
Db 590 GTGTGCGCGAGAGCG 647

RESULT 14
US-08-468-847B-1
; Sequence 1, Application US/08468847B
; Patent No. 5780263
; GENERAL INFORMATION:
; APPLICANT: Hastings, Gregg A. and Adams, Mark D.
; TITLE OF INVENTION: Human CCN-Like Growth Factor
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:

ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLIAN,
 ADDRESSEE: CECCHI, STEWART & OLSTEIN
 STREET: 6 BECKER FARM ROAD
 CITY: ROSELAND
 STATE: NEW JERSEY
 COUNTRY: USA
 ZIP: 07068
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5 INCH DISKETTE
 COMPUTER: IBM PS/2
 OPERATING SYSTEM: MS-DOS
 SOFTWARE: WORD PERFECT 5.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/468,847B
 FILING DATE: 6 June 1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: MULLINS, J.G.
 REGISTRATION NUMBER: 33,073
 REFERENCE/DOCKET NUMBER: 325800-442
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 201-994-1700
 TELEFAX: 201-994-1744
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 900 BASE PAIRS
 TYPE: NUCLEIC ACID
 STRANDEDNESS: SINGLE
 TOPOLOGY: LINEAR
 MOLECULE TYPE: CDNA
 US-08-468-847B-1

Query Match 10.6%; Score 67.2; DB 1; Length 900;
 Best Local Similarity 50.9%; Pred. No. 6.3e-07;
 Matches 274; Conservative 0; Mismatches 243; Indels 21; Gaps 4;

82 GCCTTCAGAGATGATGCGACAGAGGTATCCAGGGCTTGAGAGTACCCCGAGC---CT 138
 184 GCTTTTAAATATGATGCGACAGAAATCTTTATTCATATGTGTTAACTGTTCCAGCA 243
 139 CTCTCTGAGAACCAACGACCATGAACCGGCGGAGAGATGAGCGAGACCTCCC---CAC 195
 244 CACCCGAGCAGCAACAGCAGTGAATCAAGCCAGAAATGAGAGCGAGGATTTCACTAAC 303
 196 CATCCCTATGACCCCAAGATGTGTCCAGTACAGCTGCCGAGAGCTGCATACACCCGC 255
 304 ACTGAGCTGATGGAACACTGCGGTTCAAGTGGTTGCCGAGAACCTGCTTCAACAAA 363
 256 TTCTTCAGACAGCGCCCATGCGCAGCGCCAGCCGATCCAGATTGTTGTTGCTCCGCG 315
 364 TACATCTGTGATGCGCAGTGCACCAAGATCAGCCCTCTGAAGAGCTGTGTGTGTGAGC 423
 316 CAGTGGCGCCCGCGCGGCTGCTGCCAACGCGCATGGCGCGT-----GAAG 363
 424 GAGTGTCTTGGCCCTGCGCAGTGTCTCTTAAGTGTGAGAGGCTATGAGAACAAAGTAC 483
 364 TGGTGGCGCCCGAAGCGGATTTCCGCTGATCCCGATTCGTAACCGCGCGCAGCGG 423
 484 TGGAGCAGAGAGAGCTCCAGAGAGTGGCGGTGTGTAATGACAAAACCGTACCCAGAGA 543
 424 GTGCACTGTGTGCGCCGCGGCGCGCGCGCGCTCGCGCAGAGGTGCGTGTGTTGCGC 483
 544 ATTCAGCTGAGTGCACAGATGCGACG---ACACGCACTTACAAATATCAGATGACTACT 600
 484 TGTGCAAGTGCAGCGCTACCGCTTCCACCAACAGTGGAGTCAAGAGACTTTCGCG 543
 601 GCCTGCAAGTGCAGAGGTACACCCGCGCAGCAACAGATCCAGTACACATTGAGAGC 660
 544 CCGAGACCGCGCGCGCGAGAGGTTCCCAAGCCCGCGCGCGCGCGCGGAGGAGCCA 601

Db 661 ATGTACCTGCGACAGCGATCCAGATCAGAGAGCGGAGGAGGAGCCAGCAATCCA 718

RESULT 15
 US-09-602-877A-92/c
 Sequence 92, Application US/09602877A
 Patent No. 6432707
 GENERAL INFORMATION:
 APPLICANT: Reed, Steven G.
 APPLICANT: Xu, Jianshun
 APPLICANT: Dillon, David C.
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 FILE REFERENCE: 210121.446C5
 CURRENT APPLICATION NUMBER: US/09/602,877A
 NUMBER OF SEQ ID NOS: 107
 SOFTWARE: FastSeq for Windows Version 3.0
 SEQ ID NO 92
 LENGTH: 1692
 TYPE: DNA
 ORGANISM: Homo sapien
 US-09-602-877A-92

Query Match 10.3%; Score 65.6; DB 4; Length 1692;
 Best Local Similarity 50.7%; Pred. No. 1.6e-06;
 Matches 273; Conservative 0; Mismatches 244; Indels 21; Gaps 4;

82 GCCTTCAGAGATGATGCGACAGAGGTATCCAGGGCTTGAGAGTACCCCGAGC---CT 138
 1574 GCTTTTAAATATGATGCGACAGAAATCTTTATTCATATGTGTTAACTGTTCCAGCA 1515
 139 CTCTCTGAGAACCAACGACCATGAACCGGCGGAGAGATGAGCGACACTCCC---CAC 195
 1514 CACCCGAGCAGCAACAGCAGTGAATCAAGCCAGAAATGAGAGCGAGCATTTCACTAAC 1455
 196 CATCCCTATGACGCCCAAGATGTGTCCAGTACAGCTGCCGCGAGCTGCATACACCCGC 255
 1454 ACTGAGCTGATGGAACACTGCGGTTCAAGTGGTTGCCGAGAACTGCTTCCACAAA 1395
 256 TTCTTCAGACAGCGCCCATGCGCAGCGCCAGCCGATCCAGATTGTTGTTGCTCCGCG 315
 1394 TACATCTGTGATGCGCAGTGCACAGCATCAGCCCTTGAAGAGGCTGTGTGTGAGC 1335
 316 CAGTGGCGCCCGCGCGGCTGCTGCCAACGCGCATGGCGCGCT-----GAAG 363
 1334 GAGTGTCTTGGCCCTGCGCAGTGTCTCTTAAGTGTGAGAGGCTATGAGAACAAAGTAC 1275
 364 TGGTGGCGCCCGAAGCGGATTTCCGCTGATCCCGATTCGTAACCGCGCGCAGCGG 423
 1274 TGGAGCAGAGAGAGCTCCAGAGAGTGGCGGTGTGTAATGACAAAACCGTACCCAGAGA 1215
 424 GTGCACTGTGTGCGCCGCGGCGCGCGCGCGCTCGCGCAGAGGTGCGTGTGTTGCGC 483
 1214 ATTCAGCTGAGTGCACAGATGCGACG---ACACGCACTTACAAATATCAGATGACTACT 1158
 484 TGTGCAAGTGCAGCGCTACCGCTTCCACCAACAGTGGAGTCAAGAGACTTTCGCG 543
 1157 GCCTGCAAGTGCAGAGGTACACCCGCGCAGCAACAGATCCAGTACCAATTGAGAGC 1098
 544 CCGAGACCGCGCGCGCGCAGAGAGGTGCGCAAGCCCGCGCGCGCGCGCGGAGGAGCCA 601
 1097 ATGTACCTGAGCAAGCTCAGCATCAGAGAGCGGAGGAGGAGGAGGAGGAGGAGGAG 1040

Search completed: March 29, 2003, 01:31:02
 Job time: 59.1204 secs


```

; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepert, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 213
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-449-218D-14

```

```

Query Match          95.2%; Score 982.5; DB 4; Length 213;
Best Local Similarity 94.2%; Pred. No. 1.1e-94;
Matches 179; Conservative 3; Mismatches 3; Indels 5; Gaps 2;

```

```

QY 1 QGMAFRNDATVPIGLGEYEPPEPP--ENNQTNRRAENGGRPHHPPYDAKVSEYSCREL 58
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
    24 QGMAFRNDATVPIGLGEYEPPEPP--ENNQTNRRAENGGRPHHPPYDAKVSEYSCREL 83
DB 59 HYTRFLTDGPCRSKAPVTELVCSGCCGPARLLPNAIGRVKWRPNPDPFCIPDPRYRAOR 118
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
    84 HYTRFVTDGPCRSKAPVTELVCSGCCGPARLLPNAIGRVKWRPNPDPFCIPDPRYRAOR 143
QY 119 VOLLCPGGAAPRSKRVRLVASCKCKRLTRFNOSLKDQFGETARPOKGRKRP--GAK 175
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
    144 VOLLCPGGAAPRSKRVRLVASCKCKRLTRFNOSLKDQFGETARPOKGRKRP--GAK 203
DB 176 ANQAELENNAY 185
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
    204 ANQAELENNAY 213
QY 176 ANQAELENNAY 185
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
    204 ANQAELENNAY 213

```

```

US-09-449-218D-10
; Sequence 10, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepert, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 213
; TYPE: PRT
; ORGANISM: Cercopithecus pygerythrus
US-09-449-218D-10

```

```

Query Match          92.2%; Score 951.5; DB 4; Length 213;
Best Local Similarity 90.0%; Pred. No. 1.8e-91;
Matches 171; Conservative 10; Mismatches 4; Indels 5; Gaps 2;
QY 1 QGMAFRNDATVPIGLGEYEPPEPP--ENNQTNRRAENGGRPHHPPYDAKVSEYSCREL 58
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
    171 QGMAFRNDATVPIGLGEYEPPEPP--ENNQTNRRAENGGRPHHPPYDAKVSEYSCREL 58

```

```

DB 24 QGMAFRNDATVPIGLGEYEPPEPP--ENNQTNRRAENGGRPHHPPYDAKVSEYSCREL 83
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 59 HYTRFLTDGPCRSKAPVTELVCSGCCGPARLLPNAIGRVKWRPNPDPFCIPDPRYRAOR 118
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 84 HYTRFVTDGPCRSKAPVTELVCSGCCGPARLLPNAIGRVKWRPNPDPFCIPDPRYRAOR 143
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 119 VOLLCPGGAAPRSKRVRLVASCKCKRLTRFNOSLKDQFGETARPOKGRKRP--GAK 175
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 144 VOLLCPGGAAPRSKRVRLVASCKCKRLTRFNOSLKDQFGETARPOKGRKRP--GAK 203
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 176 ANQAELENNAY 185
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 204 ANQAELENNAY 213
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

RESULT 4
US-09-449-218D-2
; Sequence 2, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepert, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 213
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-449-218D-2

```

```

Query Match          90.4%; Score 932.5; DB 4; Length 213;
Best Local Similarity 88.4%; Pred. No. 1.7e-89;
Matches 168; Conservative 10; Mismatches 8; Indels 8; Gaps 4;
QY 1 QGMAFRNDATVPIGLGEYEPPEPP--ENNQTNRRAENGGRPHHPPYDAKVSEYSCREL 58
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
    24 QGMAFRNDATVPIGLGEYEPPEPP--ENNQTNRRAENGGRPHHPPYDAKVSEYSCREL 83
DB 59 HYTRFLTDGPCRSKAPVTELVCSGCCGPARLLPNAIGRVKWRPNPDPFCIPDPRYRAOR 118
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
    84 HYTRFVTDGPCRSKAPVTELVCSGCCGPARLLPNAIGRVKWRPNPDPFCIPDPRYRAOR 143
QY 119 VOLLCPGGAAPRSKRVRLVASCKCKRLTRFNOSLKDQFGETARPOKGRKRP--GAK 175
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
    144 VOLLCPGGAAPRSKRVRLVASCKCKRLTRFNOSLKDQFGETARPOKGRKRP--GAK 203
DB 176 ANQAELENNAY 185
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
    204 ANQAELENNAY 213

```

```

US-09-449-218D-6
; Sequence 6, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepert, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
US-09-449-218D-6

```

```

Query Match          92.2%; Score 951.5; DB 4; Length 213;
Best Local Similarity 90.0%; Pred. No. 1.8e-91;
Matches 171; Conservative 10; Mismatches 4; Indels 5; Gaps 2;
QY 1 QGMAFRNDATVPIGLGEYEPPEPP--ENNQTNRRAENGGRPHHPPYDAKVSEYSCREL 58
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
    171 QGMAFRNDATVPIGLGEYEPPEPP--ENNQTNRRAENGGRPHHPPYDAKVSEYSCREL 58

```

1 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
2 FILE REFERENCE: 240083.508
3 CURRENT APPLICATION NUMBER: US/09/449,218D
4 NUMBER OF SEQ ID NOS: 45
5 SOFTWARE: FastSeq for Windows Version 3.0
6 SEQ ID NO 6
7 LENGTH: 213
8 TYPE: PRT
9 ORGANISM: Homo sapien
10 US-09-449-218D-6

Query Match 89.5%; Score 923.5; DB 4; Length 213;
Best Local Similarity 87.9%; Pred. No. 1.5e-88;
Matches 167; Conservative 10; Mismatches 8; Indels 5; Gaps 2;

1 QGWAQFNDATFVIGLGEYEPPEPP--ENNQTMAKENGGRPHHPYDAKDVSEYSCREL 58
2 QGWAQFNDATFVIGLGEYEPPEPPLENNKTMAKENGGRPHHPETKDVSEYSCREL 83
3
4
5 59 HYTFPLTDGPRSAKPYTELVCSCGCGPARLLPNAIGRVKWRNGDFRCIPRYAOR 118
6 84 HFTRYVTDGPRSAKPYTELVCSCGCGPARLLPNAIGRVKWRNGDFRCIPRYAOR 143
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
21

MEDIUM TYPE: 3.5 INCH DISKETTE
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WORD PERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/468,847B
FILING DATE: 6 June 1995
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: MULLINS, J.G.
REGISTRATION NUMBER: 33,073
REFERENCE/DOCKET NUMBER: 325800-442
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 206 AMINO ACIDS
TYPE: AMINO ACID
STRANDEDNESS:
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
US-08-468-847B-20

Query Match 31.8%; Score 328; DB 1; Length 206;
Best Local Similarity 39.9%; Pred. No. 1,4e-26;
Matches 73; Conservative 36; Mismatches 62; Indels 12; Gaps 7;

QY 5 AFRDATEVIRGAGEYDEPP-PENNOMTMRANGRRPHHRYDAKVS---EYSCRELHY 60
DB 23 AFKDATEIIRYSHVYKVPAPHSNSTLNQNRNGR--HRSNTGLDNRTRYVQGCRLRS 80
QY 61 TRFLTDGSAKPVTELVSGGCGPARLLPNAIG---RYKWM-RPNQDPFCIPDEYRA 116
DB 81 TKYISDQCSISPLKELVCGECLPLVPLPMWIGGYGTYKYMRRSSQPMRCVNDKTRT 140
QY 117 QRVOLLPFGAAPRSKRVLYASCKCKRLTRFNQSELKDFGPEPTAP-QKGRPRGAK 175
DB 141 QRIQLQCDQDST-RTYKITVTACKCKRYTRQNHSSHNFSMSPAKVOHHRKRASK 199
QY 176 ANQ 178
DB 200 SSK 202

RESULT 12
US-09-040-229B-6
Sequence 6, Application US/09040229B
Patent No. 6432410
GENERAL INFORMATION:
APPLICANT: Harland, Richard
Hsu, David
TITLE OF INVENTION: Morphogenic Proteins
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 75 DENISE DRIVE
CITY: HILLSBOROUGH
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94010
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/040,229B
FILING DATE: 13-Mar-1998
CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:
NAME: OSMAN, RICHARD A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: B97-020-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 343-4341
TELEFAX: (650) 343-4342
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 184 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-040-229B-6

Query Match 8.8%; Score 90.5; DB 4; Length 184;
Best Local Similarity 22.9%; Pred. No. 0.063;
Matches 35; Conservative 23; Mismatches 60; Indels 35; Gaps 8;

QY 18 GEYEP-----PENNOMTMRANGRRPHH---PYDAKDVESCRELHYT--RFLTDG 67
DB 33 GAIPDPDKQPNDSQOMOTQOQSSRHRERKGTSMPEBEVLESSQALHTERKYLKD 92
QY 68 PCR-----SAKPVTELVSGGCGPARLLPNAIGRVKWMRPNQDPFR---CI 110
DB 93 WCKIQPLKQTHBEGNSNTIINRFQYGCN-SFYIPRHV-----RKREGSFQSCSPCK 145
QY 111 PDYRAQRYVOLLCPGGAAPRSRK-VLYVASCCK 142
DB 146 PKKFTTMTVTLNCPQLPQPRKKRITRYVECRIC 178

RESULT 13
US-09-040-229B-4
Sequence 4, Application US/09040229B
Patent No. 6432410
GENERAL INFORMATION:
APPLICANT: Harland, Richard
Hsu, David
TITLE OF INVENTION: Morphogenic Proteins
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 75 DENISE DRIVE
CITY: HILLSBOROUGH
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94010
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/040,229B
FILING DATE: 13-Mar-1998
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: OSMAN, RICHARD A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: B97-020-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 343-4341
TELEFAX: (650) 343-4342
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 182 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-040-229B-4

Query Match 8.6%; Score 88.5; DB 4; Length 184;

Best Local Similarity 23.4%; Pred. No. 0.1; Matches 36; Conservative 20; Mismatches 54; Indels 43; Gaps 8;

QY 23 PPEPNNQTMRAENG-----GRPPHHPYDAKDVSEYSCRELHYT--RFLTDGPCR-70
 DB 36 PPDKAQHNDSEQTSPQPGSRNRGRGGRGTAMP--GEEVLSSQGEALHVTERTYKLR 94
 QY 71 -----SAKPYTELVCSGGCGPARLLPNAIGRVKWMRPNPGRDPR---CIPDRY 114
 DB 95 QPLKQTHEDGNSRTIINRFVCGQCN-SFYIPRHIRE-----GSFQSCSFCKPKR 147
 QY 115 RAQRVOLLCPGGAAP-RSRKRVLVASCKC 142
 DB 148 TTMVVTIINCPBLQPPTKKKRVTRVKQCRC 176

RESULT 14

US-09-449-218D-42
 ; Sequence 42, Application US/09449218D
 ; Patent No. 6395511
 ; GENERAL INFORMATION:
 ; APPLICANT: Brunkow, Mary E.
 ; APPLICANT: Galas, David J.
 ; APPLICANT: Kovacevich, Brian
 ; APPLICANT: Mulligan, John T.
 ; APPLICANT: Paepert, Bryan W.
 ; APPLICANT: Van Ness, Jeffrey
 ; APPLICANT: Winkler, David G.
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
 ; FILE REFERENCE: 240083.508
 ; CURRENT APPLICATION NUMBER: US/09/449,218D
 ; CURRENT FILING DATE: 1999-11-24
 ; NUMBER OF SEQ ID NOS: 45
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 42
 ; LENGTH: 184
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-449-218D-42

Query Match 8.6%; Score 88.5; DB 4; Length 184;

Best Local Similarity 23.4%; Pred. No. 0.1; Matches 36; Conservative 20; Mismatches 54; Indels 43; Gaps 8;

QY 23 PPEPNNQ-----TMNRAENGRRPHHPYDAKDVSEYSCRELHYT--RFLTD 66
 DB 34 PPDKAQHNDSEQTSPQPGSRNRGRGGRGTAMP--GEEVLSSQGEALHVTERTYKLR 91
 QY 67 GPCR-----SAKPYTELVCSGGCGPARLLPNAIGRVKWMRPNPGRDPR---C 109
 DB 92 DWCKTQPLKQTHEDGNSRTIINRFVCGQCN-SFYIPRHIRE-----RKEGSGFQSCSFC 144
 QY 110 IPDRYRAQRVOLLCPGGAAP-RSRKRVLVASCKC 142
 DB 145 KPCKFTTMVVTIINCPBLQPPTKKKRVTRVKQCRC 176

RESULT 15

US-09-040-229B-2
 ; Sequence 2, Application US/09040229B
 ; Patent No. 6432410
 ; GENERAL INFORMATION:
 ; APPLICANT: Harland, Richard
 ; Hsu, David
 ; TITLE OF INVENTION: Morphogenic Proteins
 ; NUMBER OF SEQUENCES: 13
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
 ; STREET: 75 DENISE DRIVE
 ; CITY: HILLSBOROUGH

STATE: CALIFORNIA

COUNTRY: USA

ZIP: 94010

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/040,229B

FILING DATE: 13-Mar-1998

CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: OSMAN, RICHARD A.

REGISTRATION NUMBER: 36,627

REFERENCE/DOCKET NUMBER: B97-020-2

TELECOMMUNICATION INFORMATION:

TELEPHONE: (650) 343-4341

TELEFAX: (650) 343-4342

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 184 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-09-040-229B-2

Query Match 8.6%; Score 88.5; DB 4; Length 184;

Best Local Similarity 23.4%; Pred. No. 0.1; Matches 36; Conservative 20; Mismatches 54; Indels 43; Gaps 8;

QY 23 PPEPNNQ-----TMNRAENGRRPHHPYDAKDVSEYSCRELHYT--RFLTD 66
 DB 34 PPDKAQHNDSEQTSPQPGSRNRGRGGRGTAMP--GEEVLSSQGEALHVTERTYKLR 91
 QY 67 GPCR-----SAKPYTELVCSGGCGPARLLPNAIGRVKWMRPNPGRDPR---C 109
 DB 92 DWCKTQPLKQTHEDGNSRTIINRFVCGQCN-SFYIPRHIRE-----RKEGSGFQSCSFC 144
 QY 110 IPDRYRAQRVOLLCPGGAAP-RSRKRVLVASCKC 142
 DB 145 KPCKFTTMVVTIINCPBLQPPTKKKRVTRVKQCRC 176

Search completed: March 28, 2003, 14:20:15
 Job time: 11.3467 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 28, 2003, 17:25:48 ; Search time 2461.99 Seconds

(without alignments)
8972.025 Million cell updates/sec

Title: US-09-867-274-1

Perfect score: 759
Sequence: 1 tacgtgaagtcgctgccc.....ctgtcctctgcgcgtgttt 759

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

GenBankl: *
1: gb_ba: *
2: gb_hg: *
3: gb_in: *
4: gb_om: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sts: *
12: gb_sy: *
13: gb_un: *
14: gb_vl: *
15: em_ba: *
16: em_fun: *
17: em_hum: *
18: em_in: *
19: em_mu: *
20: em_om: *
21: em_or: *
22: em_ov: *
23: em_pat: *
24: em_ph: *
25: em_pl: *
26: em_ro: *
27: em_sts: *
28: em_un: *
29: em_vl: *
30: em_hg_hum: *
31: em_hg_inv: *
32: em_hg_other: *
33: em_hg_mus: *
34: em_hg_pln: *
35: em_hg_rnd: *
36: em_hg_mam: *
37: em_hg_vrt: *
38: em_sy: *
39: em_hgo_hum: *
40: em_hgo_mus: *
41: em_hgo_other: *

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	759	100.0	759	6 AX323453	AX323453 Sequence
2	759	100.0	2296	6 AF331844	AF331844 Homo sapi
3	759	100.0	2323	9 AF326739	AF326739 Homo sapi
4	757	99.7	2329	6 AX056687	AX056687 Sequence
5	735	96.8	2271	6 AX342535	AX342535 Sequence
6	613.2	80.8	642	9 AF326742	AF326742 Cercopit
7	536.2	70.6	674	10 AF326741	AF326741 Rattus no
8	516.8	68.1	7099	6 AX342537	AX342537 Sequence
9	516.8	68.1	21501	9 AF326736	AF326736 Homo sapi
10	516.8	68.1	80117	9 AC055813	AC055813 Homo sapi
11	516.8	68.1	94752	9 AC003098	AC003098 Homo sapi
12	502	66.1	638	10 AF326740	AF326740 Mus muscu
13	501.6	66.1	532	6 AX323455	AX323455 Sequence
14	467.8	61.6	532	4 AF326738	AF326738 Bos tauru
15	417	54.9	177744	2 AC073954	AC073954 Homo sapi
16	378	49.8	101804	2 AC098160	AC098160 Rattus no
17	378	49.8	104898	2 AC121721	AC121721 Rattus no
18	372.2	49.0	81806	10 AF326737	AF326737 Mus muscu
19	372.2	49.0	110000	2 AC068782	AC068782 Continuati (3 of
20	372.2	49.0	205277	2 AC012296	AC012296 Mus muscu
21	372.2	49.0	208135	10 AC068807	AC068807 Mus muscu
22	370.6	48.8	198508	10 AL591145	AL591145 Mouse DNA
23	158	20.8	51575	2 AC023810	AC023810 Mus muscu
24	84.2	11.1	968	10 BC021458	BC021458 Mus muscu
25	81	10.7	93790	9 AF397423	AF397423 Homo sapi
26	79.6	10.5	220967	2 AC124776	AC124776 Mus muscu
27	76.4	10.1	182253	2 AC117353	AC117353 Rattus no
28	69.8	9.2	132884	2 AC128403	AC128403 Rattus no
29	69.6	9.2	618	6 AX429979	AX429979 Sequence
30	69.6	9.2	621	6 AX429978	AX429978 Sequence
31	69.6	9.2	900	6 AR017832	AR017832 Sequence
32	69.6	9.2	1091	6 AX454484	AX454484 Sequence
33	69.6	9.2	1091	6 AX490962	AX490962 Sequence
34	69.6	9.2	1692	6 AX429953	AX429953 Sequence
35	69.6	9.2	1723	6 BC008484	BC008484 Homo sapi
36	69	9.1	154932	2 AC121167	AC121167 Rattus no
37	67.8	8.9	110958	2 AC118521	AC118521 Rattus no
38	67.6	8.9	32748	1 AB070951	AB070951 Streptomy
39	65.8	8.7	125020	9 AF429315	AF429315 Homo sapi
40	65.4	8.6	159563	2 AC112892	AC112892 Rattus no
41	65.2	8.6	1143	11 PM12A11G	AL684286 Penicillia
42	64.6	8.5	47519	2 AC110100	AC110100 Rattus no
43	64.4	8.5	2409	9 AK093408	AK093408 Homo sapi
44	64.4	8.5	125372	9 AC079155	AC079155 Homo sapi
45	64	8.4	125020	9 AF429315	AF429315 Homo sapi

ALIGNMENTS

RESULT 1
AX323453
LOCUS AX323453 759 bp DNA linear PAT 07-JAN-2002
DEFINITION Sequence 1 from Patent WO0192308.
ACCESSION AX323453
VERSION AX323453.1 GI:18094216
KEYWORDS
SOURCE
ORGANISM human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Paszty, C.J. and Gao, Y.
TITLE Cysteine-knot polypeptides: cloaked-2 molecules and uses thereof
JOURNAL Patent: WO 0192308-A 1 06-DEC-2001;

Amgen, Inc. (US)
 FEATURES Location/Qualifiers
 source 1..759
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 BASE COUNT 125 a 282 c 244 g 108 t
 ORIGIN
 Query Match 100.0%; Score 759; DB 6; Length 759;
 Best Local Similarity 100.0%; Pred. No. 1.8e-106;
 Matches 759; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 TACTGGAAGTGGGCGTCCCTCTCTGCTGTAACATGACAGCTCCACCTGCTGT 60
 Db 1 TACTGGAAGTGGGCGTCCCTCTCTGCTGTAACATGACAGCTCCACCTGCTGT 60
 QY 61 CTGCTCTGCTGCTGTAACACAGCTTCCGTGTAGTGAAGGCGGAGGCTGCAAGCG 120
 Db 61 CTGCTCTGCTGCTGTAACACAGCTTCCGTGTAGTGAAGGCGGAGGCTGCAAGCG 120
 QY 121 TTCAAGATGATGCGACGAAATCATCCCGAGCTGGAGATACCCCGAGCTCCACCG 180
 Db 121 TTCAAGATGATGCGACGAAATCATCCCGAGCTGGAGATACCCCGAGCTCCACCG 180
 QY 181 GAGCTGGAACAACAAGACCATGAAACCGGCGGAGAACGAGGCGGCTCCCAACA 240
 Db 181 GAGCTGGAACAACAAGACCATGAAACCGGCGGAGAACGAGGCGGCTCCCAACA 240
 QY 241 CCTTTGAGACCAAAAGCGTGTCCGAGTACAGCTGCCGAGCTGCACTTACCCGCTAC 300
 Db 241 CCTTTGAGACCAAAAGCGTGTCCGAGTACAGCTGCCGAGCTGCACTTACCCGCTAC 300
 QY 301 GTGACGATGGGCGGCGGAGCGGCAAGCGGCTGCAAGCTGCTGTCCGCGCAG 360
 Db 301 GTGACGATGGGCGGCGGAGCGGCAAGCGGCTGCAAGCTGCTGTCCGCGCAG 360
 QY 361 TGGCGGCGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
 Db 361 TGGCGGCGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
 QY 421 GGGCGGCGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
 Db 421 GGGCGGCGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
 QY 481 CCGGCTGATGAGGCGGCGGCGGCGGCAAGGCTGCTGCTGCTGCTGCTGCTGCT 540
 Db 481 CCGGCTGATGAGGCGGCGGCGGCGGCAAGGCTGCTGCTGCTGCTGCTGCTGCT 540
 QY 541 CGCCTCAACCGCTTCAACAACAAGCTCAAGGCTTCAAGGCTTCAAGGCTTCAAGG 600
 Db 541 CGCCTCAACCGCTTCAACAACAAGCTCAAGGCTTCAAGGCTTCAAGGCTTCAAGG 600
 QY 601 CCGGAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 660
 Db 601 CCGGAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 660
 QY 661 CTGGAGAACGCTACTAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 720
 Db 661 CTGGAGAACGCTACTAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 720
 QY 721 GAACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 759
 Db 721 GAACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 759

RESULT 2
 AF331844 2296 bp mRNA linear PRI 06-MAR-2001
 LOCUS AF331844
 DEFINITION Homo sapiens SOST (SOST) mRNA, complete cds.
 ACCESSION AF331844
 VERSION AF331844.1 GI:13236417
 KEYWORDS
 SOURCE Homo sapiens.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 2296)
 AUTHORS Balenans, W., Ebeling, M., Patel, N., van Hul, E., Olson, P.,
 Droszegi, M., Laczka, C., Wuyts, W., van den Ende, J., Willems, P.,
 Paes-Alves, A.F., Hill, S., Bueno, M., Ramos, F.J., Tacon, P.,
 Dikkers, F.G., Stratakis, C., Lindpaintner, K., Vickery, B.,
 Foerzler, D. and Van Hul, W.
 TITLE Increased bone density in sclerosteosis is due to the deficiency of
 a novel secreted protein (SOST)
 JOURNAL Hum. Mol. Genet. 10 (5), 537-543 (2001)
 MEDLINE 21096930
 PUBMED 1181578
 REFERENCE 2 (bases 1 to 2296)
 AUTHORS Balenans, W., Ebeling, M., Patel, N., Vickery, B., Foerzler, D. and Van
 Hul, W.
 TITLE Direct Submission
 JOURNAL Submitted (22-DEC-2000) Medical Genetics, University of Antwerp,
 Universiteitsplein 1, Antwerp B2610, Belgium
 FEATURES Location/Qualifiers
 source 1..2296
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="17"
 /map="17q12-q21"
 1..2296
 /gene="SOST"
 38..679
 /gene="SOST"
 /note="mutations may cause sclerosteosis"
 /codon_start=1
 /product="SOST"
 /protein_id="AAK16158.1"
 /db_xref="GI:13236418"
 /translation="MQLPLALCLVCLLVHTAFRVEVGQWQAFKNDATIIIPELGEYP
 EEPPEENNTNMRABNGRPPHPPETKDVSVSCRELFHTVYVDGCRSAKPYTE
 LVSCGCGPARLLPNAIGRWMPSPGDPKCIPIPRARORVOLCPGGEAPARVLR
 LVASCKKRLTRHNOSELDKFTGARPPQKGRKPRPARSAANQAELENNV"
 BASE COUNT 576 a 631 c 607 g 482 t
 ORIGIN
 Query Match 100.0%; Score 759; DB 9; Length 2296;
 Best Local Similarity 100.0%; Pred. No. 1.4e-106;
 Matches 759; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 TACTGGAAGTGGGCGTCCCTCTCTGCTGTAACATGACAGCTCCACCTGCTGT 60
 Db 2 TACTGGAAGTGGGCGTCCCTCTCTGCTGTAACATGACAGCTCCACCTGCTGT 61
 QY 61 CTGCTCTGCTGCTGTAACACAGCTTCCGTGTAGTGAAGGCGGAGGCTGCAAGCG 120
 Db 62 CTGCTCTGCTGCTGTAACACAGCTTCCGTGTAGTGAAGGCGGAGGCTGCAAGCG 121
 QY 121 TTCAAGATGATGCGACGAAATCATCCCGAGCTGGAGATACCCCGAGCTCCACCG 180
 Db 122 TTCAAGATGATGCGACGAAATCATCCCGAGCTGGAGATACCCCGAGCTCCACCG 181
 QY 181 GAGCTGGAACAACAAGACCATGAAACCGGCGGAGAACGAGGCGGCTCCCAACA 240
 Db 182 GAGCTGGAACAACAAGACCATGAAACCGGCGGAGAACGAGGCGGCTCCCAACA 241
 QY 241 CCTTTGAGACCAAAAGCGTGTCCGAGTACAGCTGCCGAGCTGCACTTACCCGCTAC 300
 Db 242 CCTTTGAGACCAAAAGCGTGTCCGAGTACAGCTGCCGAGCTGCACTTACCCGCTAC 301
 QY 301 GTGACGATGGGCGGCGGAGCGGCAAGCGGCTGCAAGCTGCTGCTGCTGCTGCTGCT 360
 Db 302 GTGACGATGGGCGGCGGAGCGGCAAGCGGCTGCAAGCTGCTGCTGCTGCTGCTGCT 361
 QY 361 TGGCGGCGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
 Db 362 TGGCGGCGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 421

QY	421	GGGGCCGACATTCCGCTGATCCCGACCGCTACCGCGAGCCGCTGACGCTGTGT	480
Db	422	GGGCCCCGACTTCCGCTGATCCCGACCGCTACCGCGAGCCGCTGACGCTGTGT	481
QY	481	CCCGGTGTGAGGCGCCGCGCGCGAGAGTGCGCTGTGCTGTGCAAGTGCAAG	540
Db	482	CCCGGTGTGAGGCGCCGCGCGCGAGAGTGCGCTGTGCTGTGCAAGTGCAAG	541
QY	541	CGCTCAACCCGCTTCCACACAGTCGGAAGTCAAGAGCTTCGGAGACCGAGCGCTCG	600
Db	542	CGCTCAACCCGCTTCCACACAGTCGGAAGTCAAGAGCTTCGGAGACCGAGCGCTCG	601
QY	601	CCGAGAAAGGCGCGAAGCGCGCGCGCGCGCGCGAGCGCGCAAGCCACAGCGAG	660
Db	602	CCGAGAAAGGCGCGAAGCGCGCGCGCGCGCGCGAGCGCGCAAGCCACAGCGAG	661
QY	661	CTGAGAAACGCTTACAGAGCCCGCGCGCGCGCTTCCACCGAGCGCGCGCGCTT	720
Db	662	CTGAGAAACGCTTACAGAGCCCGCGCGCGCGCTTCCACCGAGCGCGCGCGCTT	721
QY	721	GACCCGCGCCCAATTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	759
Db	722	GACCCGCGCCCAATTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	760
RESULT_3	AF326739		
LOCUS	AF326739	2123 bp	mRNA linear PRI 28-FEB-2001
DEFINITION	Homo sapiens sclerostin mRNA, complete cds.		
ACCESSION	AF326739		
VERSION	AF326739.1	GI:13161019	
KEYWORDS			
SOURCE	Homo sapiens.		
ORGANISM	Homo sapiens.		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 (bases 1 to 2123) Brunkow,M.E., Gardner,J.C., Van Ness,J., Paeppe,B.W., Kovacevich,B.R., Proli,S., Skonier,J.E., Zhao,J., Sabo,P.J., Fu,Y.H., Alisch,R.S., Gillett,L., Colbert,T., Tacconi,P., Galas,D., Hamersma,H., Beighton,P. and Mulligan,J.T. Bone dysplasia sclerosteosis results from loss of the SOST gene product, a novel cystine knot-containing protein Am. J. Hum. Genet. 68 (3), 577-589 (2001) 21090529		
TITLE			
JOURNAL			
MEDLINE	11179006		
REFERENCE	2 (bases 1 to 2123) Brunkow,M.E., Gardner,J.C., Van Ness,J., Paeppe,B.W., Kovacevich,B.R., Proli,S., Skonier,J.E., Zhao,J., Sabo,P.J., Fu,Y.H., Alisch,R.S., Gillett,L., Colbert,T., Tacconi,P., Galas,D., Hamersma,H., Beighton,P. and Mulligan,J.T. Direct Submission Submitted (07-DEC-2000) Genomics, Celltech Chiroscience Inc., 1631 220th St. SE, Bothell, WA 98021, USA		
JOURNAL			
FEATURES	location/Qualifiers		
source	1..2123 /organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="17" /map="17q21" 1..47 48..689 /codon_start=1 /product="sclerostin" /protein_id="AAK13454.1" /db_xref="GI:13161020" /translation="MQLPLALCVLLVHTAARVVEGSGMOAFKNDATETIIPELGEYH EPPLENNKTKMRAENGSRPHHPFEKDVSEYSCREIHTFRTYTDPCSSAKVTVT LVSCGCGPARLLPAVIGKGRKMRSPGDFRCIPRYRAORVOLLCPGEAPRAKAVN LVAASKRILTRFNHQSLEKDFGTEAARQGRKRPRAKSAKQAELNNAY"		
5'UTR			
CDS			
3'UTR			
BASE COUNT	590 a 635 c 614 g 484 t		

ORIGIN

	Query Match	Similarity	Score	DB	Length
	100.0%	100.0%	759	DB 9	2323
	Local	Similarity	100.0%	Pred. No. 1.4e-106	
	Matches	759	Conservative	0	Mismatches 0
					Indels 0
					Gaps 0

	Query	Subject	Score	DB	Length
1	TACTGGAAGGTGGCGTGCCTCTCTGAGTGTACATGCAAGCTCCCACTGGCCCTGTGT	60			
12	TACTGGAAGGTGGCGTGCCTCTCTGAGTGTACATGCAAGCTCCCACTGGCCCTGTGT	71			
61	CTGCTGCGCCGCTGCTGTACACAGAGCCCTCCGCTGTAGTGGAGGCGCAGGGGTGGCAGCG	120			
72	CTGCTGCGCCGCTGCTGTACACAGAGCCCTCCGCTGTAGTGGAGGCGCAGGGGTGGCAGCG	131			
121	TTCAAGATGATGTCACAGGAATCATCCCGAGCTCGAAGATACCCCGAGCTCCACCG	180			
132	TTCAAGATGATGTCACAGGAATCATCCCGAGCTCGAAGATACCCCGAGCTCCACCG	191			
181	GAGCTGAGGAACCAACAGATACATGACCGGGGGGAGAACGAGGGGGGGCCCTCCACACAC	240			
192	GAGCTGAGGAACCAACAGATACATGACCGGGGGGAGAACGAGGGGGGGCCCTCCACACAC	251			
241	CCCTTTGAGACCAAGAGAGCTGTCCGAGTACAGTGCAGCTGCAGCTCACTTCAACCCGCTAC	300			
252	CCCTTTGAGACCAAGAGAGCTGTCCGAGTACAGTGCAGCTGCAGCTCACTTCAACCCGCTAC	311			
301	GTGACCGATGGGCGCGTGCAGCGCCAGCCGCTCACCCAGCTGTGTGTCTCCGGCCAG	360			
312	GTGACCGATGGGCGCGTGCAGCGCCAGCCGCTCACCCAGCTGTGTGTCTCCGGCCAG	371			
361	TGCGGCGCGCGCGCGCTGCTCCCAAGCCCAAGCCGAGCGGCGGCAATGAGTGGAGCACTAGT	420			
372	TGCGGCGCGCGCGCGCTGCTCCCAAGCCCAAGCCGAGCGGCGGCAATGAGTGGAGCACTAGT	431			
421	GCGGCCGAGCTTCCGCTGTCATCCCGACCGCTACCGCGGAGCGCGTGCAGCTCTGTGT	480			
432	GCGGCCGAGCTTCCGCTGTCATCCCGACCGCTACCGCGGAGCGCGTGCAGCTCTGTGT	491			
481	CCCGGTGTGAGGCGCCCGCGCGCGGCAAGTGGCGCTGTGTGTGTGTGTGTGTGTGTGTGT	540			
492	CCCGGTGTGAGGCGCCCGCGCGGCAAGTGGCGCTGTGTGTGTGTGTGTGTGTGTGTGT	551			
541	GCGCTCACCCGCTTCCAGCAACAGTGGAGTCTCAAGAGCTTCCGAGCCGAGCGCTGTGT	600			
552	GCGCTCACCCGCTTCCAGCAACAGTGGAGTCTCAAGAGCTTCCGAGCCGAGCGCTGTGT	611			
601	CCGCAAGAGGCGCGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG	660			
612	CCGCAAGAGGCGCGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG	671			
661	CTGGAAGAGCTTACATGAGCT	720			
672	CTGGAAGAGCTTACATGAGCT	731			
721	GAAACCGCTT	759			
732	GAAACCGCTT	770			

	RESULT 4	AX056687	2329 bp	DNA	linear	PAT 17-JAN-2001
LOCUS	AX056687					
DEFINITION	Sequence 19 from Patent WO0075317.					
ACCESSION	AX056687					
VERSION	AX056687.1	GI:12309667				
KEYWORDS						
SOURCE	Human.					
ORGANISM	Homo sapiens					
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
REFERENCE	1 (bases 1 to 2329)					
AUTHORS	Botstein,D.A., Goddard,A., Gurney,A.L., Smith,V., Watermanbe,C.K. and Wood,W.I.					

